LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Department ”Institut für Informatik”
Lehr- und Forschungseinheit Medieninformatik
Prof. Dr. Heinrich Hußmann

Masterarbeit

Offline (Quantum) Key Distribution

Maximilian Wagner
Wagner.Maximilian@campus.lmu.de

Bearbeitungszeitraum:	07. 03. 2022 bis 05. 09. 2022
Betreuer:	Sarah Delgado Rodriguez, Sarah Prange
Verantw. Hochschullehrer:	Prof. Florian Alt

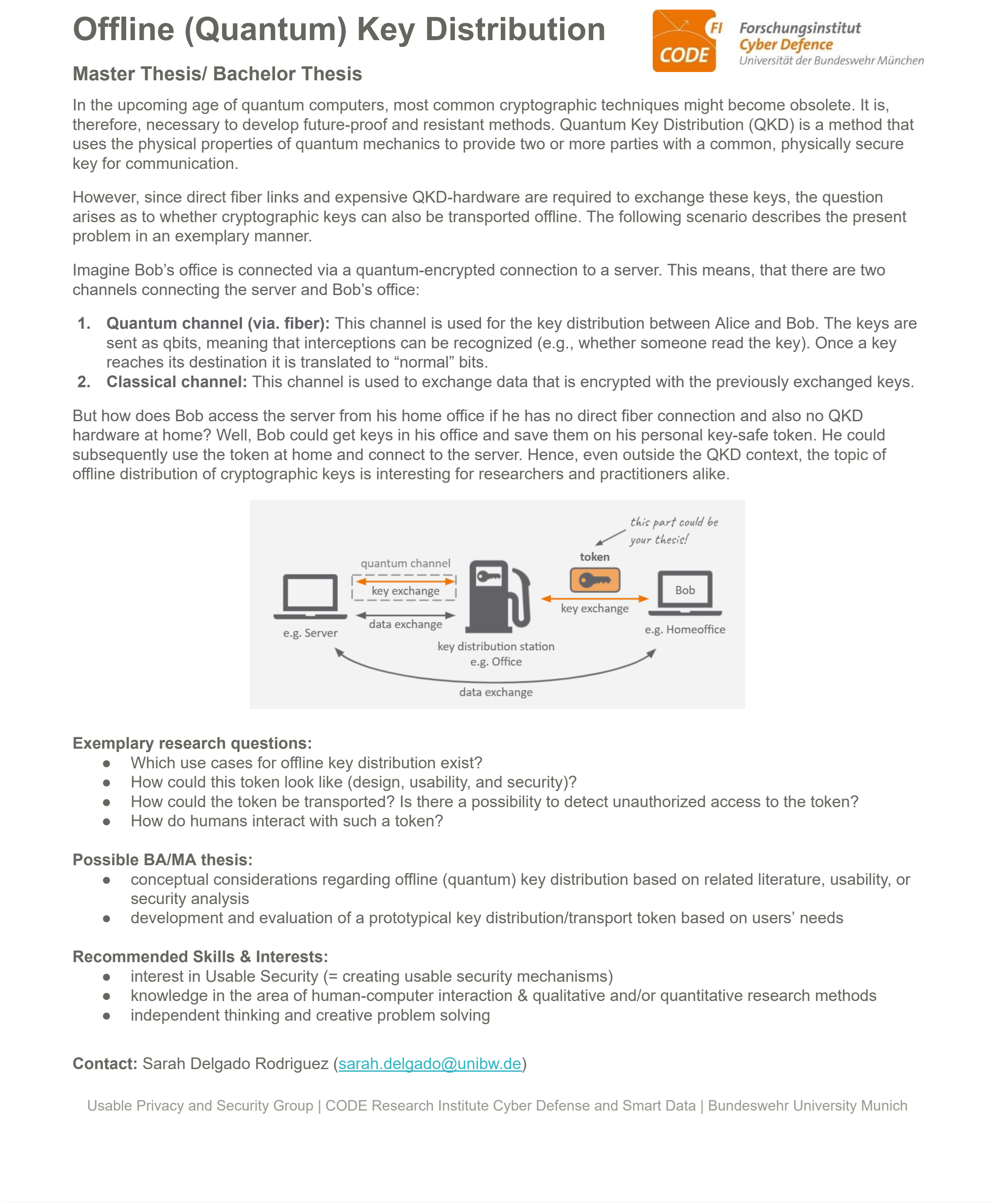
Summary

In dieser Arbeit wird ein Home Office Szenario als möglicher Anwendungsfall für offline SchlüsselaustauschVerfahren vorgestellt und für diesen dann drei mögliche Konzepte entworfen. Die Konzepte RFID-Sticker, USB-Keychain und QR-Device werden dann im Rahmen einer Nutzerstudie sowohl miteinander als auch mit den Baselines Papier und Smartcard verglichen. Im Rahmen der Studie müssen die Teilnehmer mit allen fünf Konzepten das Ein-/Auslesen und den Transport kryptographischer Schlüssel testen und diese dann nach Aspekten der Benutzerfreundlichkeit, Verständlichkeit, Vertrauen und gefühlter Sicherheit bewerten. Als Resultat ergibt sich, dass nur eines der drei Konzpte im Vergleich zur Baseline im Bezug auf Nutzerfreundlichkeit besser abschneidet, allerdings in den anderen Kategorien schlechter. Gesamt schneidet die aktuell verwendete Standardsmartcard am besten ab, da sie dem RFID-Sticker in den Kategorien Vertrauen und gefühlte Sicherheit überlegen ist.

Abstract

In this thesis a home office scenario is developed as a possible use case for offline key distribution. Three concepts are presented as possible solutions for this use case. The concepts of RFID sticker, USB-keychain and QR-Device are evaluated with a user study and compared to two baselines, paper and smartcards. During the user study the participants have to evaluate the user interactions for reading, writing and transporting cryptographic keys in regards to usability, tangibility, trust and perceived security. The results show, that only the RFID sticker of the three concepts surpasses the better baseline in usability, but scores lower in the other categories. The smartcard baseline generally performs best, as it scores higher in all other categories than the smartcard.

Topic



Offline (Quantum) Key Distribution

Forschungsinstitut Cyber Defence
Universität der Bundeswehr München

Master Thesis/ Bachelor Thesis

In the upcoming age of quantum computers, most common cryptographic techniques might become obsolete. It is, therefore, necessary to develop future-proof and resistant methods. Quantum Key Distribution (QKD) is a method that uses the physical properties of quantum mechanics to provide two or more parties with a common, physically secure key for communication.

However, since direct fiber links and expensive QKD-hardware are required to exchange these keys, the question arises as to whether cryptographic keys can also be transported offline. The following scenario describes the present problem in an exemplary manner.

Imagine Bob's office is connected via a quantum-encrypted connection to a server. This means, that there are two channels connecting the server and Bob's office:

1. **Quantum channel (via. fiber):** This channel is used for the key distribution between Alice and Bob. The keys are sent as qbits, meaning that interceptions can be recognized (e.g., whether someone read the key). Once a key reaches its destination it is translated to "normal" bits.
2. **Classical channel:** This channel is used to exchange data that is encrypted with the previously exchanged keys.

But how does Bob access the server from his home office if he has no direct fiber connection and also no QKD hardware at home? Well, Bob could get keys in his office and save them on his personal key-safe token. He could subsequently use the token at home and connect to the server. Hence, even outside the QKD context, the topic of offline distribution of cryptographic keys is interesting for researchers and practitioners alike.

Exemplary research questions:

- Which use cases for offline key distribution exist?
- How could this token look like (design, usability, and security)?
- How could the token be transported? Is there a possibility to detect unauthorized access to the token?
- How do humans interact with such a token?

Possible BA/MA thesis:

- conceptual considerations regarding offline (quantum) key distribution based on related literature, usability, or security analysis
- development and evaluation of a prototypical key distribution/transport token based on users' needs

Recommended Skills & Interests:

- interest in Usable Security (= creating usable security mechanisms)
- knowledge in the area of human-computer interaction & qualitative and/or quantitative research methods
- independent thinking and creative problem solving

Contact: Sarah Delgado Rodriguez (sarah.delgado@unibw.de)

Usable Privacy and Security Group | CODE Research Institute Cyber Defense and Smart Data | Bundeswehr University Munich

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, September 5, 2022

Wagner Max
....

Contents

1 Introduction

This thesis focuses on conception and evaluation of different physical tokens or hardware devices that can be used or are currently used to exchange keys offline. While having 'quantum' in the title, the content presented here is not limited to quantum key distribution or general key distribution, but can generally be applied to the transport of digital secrets. The first subsection provides a short overview of the motivation behind this thesis, namely the issue what happens when a cryptographic key exchanged via a quantum secure channel has to be transmitted over a non quantum secure channel. The next subsection gives a short description about the contributions this thesis makes towards the usability of specific devices for the offline exchange of cryptographic keys.

1.1 Motivation

With the advancement of quantum communication in the last few years, more research is done in that field. One important aspect in the field of quantum computing is quantum key distribution.

Quantum key distribution allows the exchange of cryptographic keys over a quantum secure channel. Without diving to deep into the aspects of quantum communication, this allows the communicating entities to detect (with a probability of 50% per transmitted bit), if the message has been eavesdropped on.

If an intercept has been detected a new set of keys can be exchanged. While this explanation is extremely simplified, it is sufficient enough to say, that a key exchanged via a quantum secure channel is secure and only known to the two communicating entities.

What happens now, if the key has to be transmitted through a non-quantum secure communications channel? Let us assume, that an attacker, capable of intercepting (and trying to decode) a key exchanged on a quantum secure channel, also has the capabilities of intercepting the transmission on the other, classic channel.

While most communication on that channel will still be encrypted, that encryption (or rather the key exchange prior to encryption) will probably be based on current key exchange schemes. Those based on public key cryptography (and the difficulty of factoring large primes) can be assumed to be computationally secure.

What if a level of physical security for the transmission of the keys is desired that can match the level of security of a quantum secure channel?

The answer could be the transport of the keys through offline, out-of-band channels, that increase the effort required for an attacker to gain access to the keys.

If the keys were to be transported offline, one could assume that there are dedicated devices for the offline key exchange, three potential concepts will be presented in this thesis.

1.2 Contribution

The contribution of this thesis to the topic of offline (quantum) key distribution as well as the research questions will be provided in this section. The word quantum in the thesis title is set in parentheses, which was not a mistake. It was done on purpose, as the only relevant connection to the field of quantum key distribution this thesis has, is rooted in the motivation section and the original problem description. Apart from the motivation, the fact, that in the original problem, a quantum key should be exchanged using an offline exchange method does not really matter. Why is that? Because from the point on where the keys have been determined on both endpoints of a quantum communication channel, for all practical purposes the key that then has to be transmitted onto the offline key exchange device, is just a sequence of bits, which can be treated like any other symmetric key.

With that in mind, in this thesis a simple use case for the use of offline key exchange devices in an enterprise environment is developed. The use case is then presented to a focus group to gather insights and ideas on how the device could work or what it could look like. Three prototypes are created from the focus group suggestions and compared and evaluated for aspects of usability, tangibility and perceived trust and security with the use of a user study.

This directly leads to the three research questions this thesis wants to answer:

- Q1: How do users perceive the usability of different key exchange devices (compared to a baseline)?
- Q2: How do the differences in tangibility affect the user experience?
- Q3: How do the different hardware devices effect the perception of trust and security in the devices?

2 Related Work

This chapter provides a overview of already existing methods of key transportation devices as well as related devices. This chapter may be rather short, as generally there is not much relevant related work, or research at all, in the area of offline key distribution. Many papers on key exchange schemes mention the use of trusted couriers as alternative [27], [37] for secure key exchanges, but searching for the term ”trusted courier” in the ACM Guide to Computing Literature brings up only 30 results, the term ”offline key exchange” 3 results and ”offline key distribution” 12 results. Even though searching in the IEEEXplore database yields more results (”trusted courier”: 5, ”offline key exchange”: 89, ”offline key distribution”: 221), Most of those results do not have any relevancy for this thesis as most of them focus on various technical aspects but not on aspects mentioned in this thesis research questions.

2.1 Hardware

Since the goal of this thesis is the creation and evaluation of different types of hardware tokens for offline key exchange, this section takes a look at different types of hardware authentication tokens and similar devices, as they can be seen as having some similarities in purpose and usage.

The first set of tokens are various two factor authentication tokens, with one of the most common USB based ones being Yubikeys [1].

Those and other web authentication systems have been researched in regards to their usability and use acceptance.

Reese et. al compared the usability of hardware 2FA tokens with SMS or push notification based systems [34]. Problems with the hardware based system did not occur directly while using the devices but were caused by the setup process. Acemyan et al. report a high rate of successfull logins using usb security keys in comparison to an app or SMS based solution [1], Ciolino et al. mentions a few common problems appearing with the YubiKey [10]. Similar problems in the usability of YubiKeys were reported by Das et al., as users were confused by the form factor and design of the usb key, thinking that a simple button was instead a capacitive fingerprint sensor [12]. There are other types of dedicated single purpose hardware devices, that can be used as a second factor, some of them were studied by Krol et al.[22] While not being usb based, users reported a low level of satisfaction in using those other hardware based authentication systems, as they had sometimes still to provide multiple credentials to successfully login or still remember information in addition to their passwords. Users wished for an authentication system using biometrics as well as reducing cognitive and physical effort, as they don’t need to remember addition information or carry an additional device with them. Having to carry an additional device with them was also mentioned in [17], where YubiKeys were used in passwordless authentication schemes. A bigger

[1] https://www.yubico.com/products/security-key/

factor for not using the YubiKey as a single authentication mechanism was a lack of trust in the device, as the participants were missing a mental model for authentication without a password.

USB based devices are not the only similar device, as smartcards in comparison are already widely used to store private keys for signing and decryption purposes. An evaluation of the usability of smartcards compared to USB based tokens can be found in [11]. In the experiment conducted, smartcards achieve lower ratings in usability and mobility than the USB devices compared against. The authors mention, that this was caused by the fact, that for the smartcard an additional reader had to be carried, plugged in and installed at each locations, leading to problems when participants forgot the reader at the last location. The most relevant aspects here include a measurement of what types of errors occurred when using smartcards, with the most common type of error being the card inserted the wrong way or inserting the card not completely. Morse et. al looked at smartcard usability and reported a overall high level of satisfaction with smartcard use, when replacing the username/password combination for the login workflow. Even though some issues with card readers existed, the use of a PIN and smartcard was seen as favourable, with most participants willing to recommend smartcards to their coworkers [28]. The usability ans security of smartcard and other authentication mechanisms was evaluated by Braz and Robert, with high ratings for security and average ratings for usability [8].

These are not the only type of authentication tokens that exist, as Stajano explores an interesting concept of a device, that is only unlocked and able to authenticate its user, when a minimum amount of smaller sibling devices is in close proximity [42]

2.2 Transportation

As mentioned above, when talking about offline key exchange, the term of trusted courier comes up. Couriers have been used for thousands of years, and while they are still a popular choice for the transport of sensitive information, the enforcement of security measures is costly and not practical for larger systems [13][7]. Alléaume et. al mention the differences as well as the common features between quantum key distribution and trusted courier key exchange. The aspects mentioned here in regard to trusted couriers consist of trust, reliability and bandwidth. The problem with trust in a trusted courier situation is, that corruption of the courier can not be detected, for reliability, that scaling a key exchange scheme based on trusted couriers becomes difficult and expensive [2]. With that in mind, when talking about the use of trusted couriers in the context of this thesis, we follow the general practice of assuming perfect security for a key exchange via trusted courier.

2.3 Data Transfer

This section presents some works on data transmission methods through human interaction There are many papers on different data transfer methods, most focus on technical aspects of data transfer. As we are specifically interested in how data can be transferred to and from the hardware token. The simplest method would be typing the key, either from the token into the target system, or directly into the token. The field of typing has been intensively researched. Rumelhart et. al found that having to type random sequences of character is slower and more error prone, than typing even an text in a foreign language [36]. Typing speed is related to the keyboard layout, but even on small mobile device keyboards, people do not type much slower than on regular keyboards [30]. Yang et. al found out, that text entry methods do affect password strength, due to special characters being harder to reach on mobile keyboards [44]

An other method could be the use of speaking PINs or passwords a concept presented by Bailey et. al [4]. While this focuses on authentication mechanisms on AR glasses, it would allow the transmission of a key from a token to the target device, when reading the key out loud. Automatic speech recognition systems still face challenges, but is adequate for smaller vocabularies [29], which would entail the character set commonly used for password.

A third method could be key transfer through the use of the files system. The usability of files systems and storing and retrieving files has been researched. Barreau and Nardi found that users had not problem retrieving their work information, as it is commonly used and remembered where it is stored [6]. Even though the usability of operating systems generally has improved over the last years [31], there are reports, anecdotal at best, that younger generations may not be able to navigate hierarchical files systems anymore. It is speculated, this may be caused through various modern (cloud) storage solution and user interfaces abstracting away the hierarchical file system structure and the users depending on well working search functions [45].

2.4 Conclusion

We found, that while there exist various devices already for the transport of cryptographic keys, the amount of works in the specific case of measuring human factors in offline key exchanges is practically non existent. While this makes the design of a dedicated hardware key transport device more difficult, as only some inspiration can be gained through the above mentioned works, we try to design or select the devices for this thesis in a comprehensible way.

3 Concepts

In this chapter the concept ideas for the hardware token are developed. As only some inspiration can be gained from related work, this allows us to completely start from scratch in what could be used for the key transport device. This chapter gives an overview of the design process and thoughts behind the final implemented concepts.

3.1 Developing a use case

Before concepts and interactions of the hardware tokens can be formulated, a use case for the hardware tokens has to be specified, as different aspects of the hardware token could depend on external factors.

In this section the general use case for the hardware token will be described. The use case will be based on a common and simple set of interactions of a user with the hardware token in the key exchange process.

Firstly, an example for the use case will be provided, with a more formalized description of the various interactions. These will be used as a basis for the functionality and usability of the token, as well as the evaluation of tokens performance.

3.1.1 Considerations for the use case setting

First the general setting, in which the use of offline key exchanges would even be a realistic necessity, has to be considered. In the next step it has to be considered, how likely it is, that for some specific kind of situation one would have to use a dedicated hardware token for the exchange of cryptographic keys.

In the setting of everyday life of individuals in the role of consumers or end-user, such a situation is highly unlikely. Many people only want their technology to work, without being interested in the exact specifics behind it. Users are already often unwilling to use security features if their use is inconvenient[3].

An other setting would be a work related e.g. enterprise setting. Here it is quite easy to imagine being forced to use some tools and technologies because it is mandated by the employer. Even if a person would not use 2FA authentication tokens in their private life, if those are required to access systems at work the employee has no choice in that matter. In this setting it could be quite likely that an employee has to use some dedicated hardware token, issued by the employer, to transport

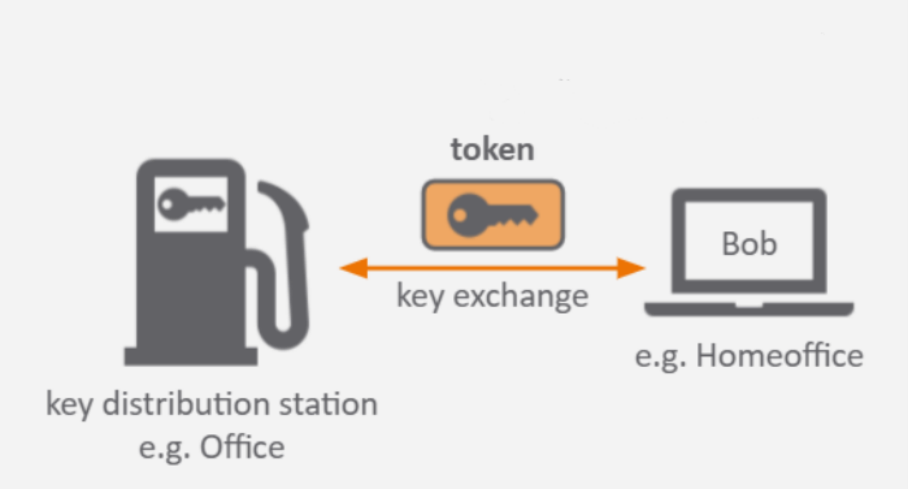


Figure 3.1: Hardware Token Use Case (image adapted from the original thesis task description, see p. c).

or store cryptographic keys or sensitive data. This is is already a common case in enterprise environments, where e.g. the private keys for decrypting emails are stored on an employees badge.

A third setting, in which the offline transport of cryptographic keys could be a real use case scenario, is a military or intelligence contexts. But due to the high level of security often required in these fields, it can be assumed that trade-offs in usability to increase the level of security are commonly made, which leads into a different direction than the proposed research questions in chap 1.2.

Taking these three ideas for settings into consideration, only setting two would allow the construction of a reasonable use case, that does not feel too artificial.

Due to the corona virus pandemic and the increase in work from home situation, this leads to a simple and relatable scenario.

3.1.2 Example Use Case: Home office

There could be different use cases in an enterprise setting. For example a key has to be exchanged between two company locations. Another use case would be if some kind of outside field equipment has to be supplied with new keys as part of a key rotation scheme, which is how key updates are distributed between components of the European train control systems [25]. Other common use cases are key distributions in IoT networks[20] or smart grid settings [43].

Compared to those two examples the home office scenario is relatable, easy to understand as many people have already experience the general setting, and it should be not too complicated to simulate for a user study. This use case scenario has an additional advantage, namely that this scenario works without a trusted courier to transport the keys. While the employee has to physically transport the key storage device home this does not change the fact that the employee already is the intended recipient for the keys. This case is therefore an example of receiving encryption keys directly through 'personal' contact.

For the use case consider the following example (also depicted in fig 3.1): Since some time ago, Bob is working from home. While working from home has many advantages, Bob faces a new problem. Bobs employer has introduced a new system with quite a high degree of security. One problem here is, that communication with the system is only allowed when specific requirements are fulfilled. One of those requirements is, that communication from outside sources, e.g. employees working from home, is only possible using an encrypted connection. The additional caveat is, that the encryption key necessary for establishing an encrypted connection must not be sent over the internet, as this has been deemed insecure.

To facilitate the key exchange, Bob has received some kind hardware token/device, specifically designed for the transport of encryption keys. To successfully work from home with the new system, a key has to be generated by a key distribution station and transferred onto the device which in turn has to be transported to Bob's home. Bob can then use the key to connect to the new system.

Under the assumption the user has a specific device for the transport of the key, which allows the storage and retrieval of the keys, they now have to perform the following steps to transport the keys from the office to their home office:

- The user requests a new key to be generated by the key distribution station
- The user stores the key on the hardware token, either through some form of direct or indirect interaction with the station and/or the token
- The user transports the token to the keys destination
- The user transfers the key to the target computer, where the key can be used, either through direct or indirect interaction with the computer or the token

The steps have been described only in a non specific manner, as the interactions can take many forms, from reading the key and manually inputting it into its target to simply presenting the device showing some kind of code to the computer, making further interaction unnecessary. The specifics of each interaction depend on the capabilities and functionalities of the hardware token itself. These will be determined in the following sections.

3.2 Focus Group

To gather insights on how the individual user interaction steps could look like, a focus group was conducted, where for each of the steps one question was created.

3.2.1 Apparatus

To gather insights about what aspects of the hardware token are important for users a small focus group was conducted. The goal of the focus group was to collect ideas on functional and usability aspects of the hardware token, as well as an evaluation on the originality and feasibility of those aspects by the possible users.

The focus group was held online using Zoom. In addition, to provide an overview over the discussion, as well as give the participants the opportunity to write down and later evaluate their ideas, the online collaboration/whiteboard tool Mural was used. This allowed us, to have separate areas for each question to preserve the relation between questions and answers.

3.2.2 Participants

The participants were selected from university students and recent graduates, with the requirement of having finished their first degree and this degree being rooted in computer science or an adjacent field. This requirement was chosen for the following reasons: By taking people from a tech background it can be reasonably assumed that the participants have been exposed to a variety of technologies that could be used as an inspiration on the design of a new hardware token, as well as the assumption that a having tech background making the participants more open to new and innovative technologies. Having finished their first degree, it is highly likely, the participants have some inherent understanding on the difficulties of (offline) key exchange schemes and the security implications. At the same time, with the participants being recent graduates, they can bring their experiences to the table on the challenges such a hardware token has to face in a corporate environment.

3.2.3 Question guide

Following the recommended focus group methods [24], a set of of open questions was created to guide though the focus group.

The questions were created with the purpose of representing the individual steps of user interaction of the use case from chapter 3.1. The order and the specific wording of questions was chosen with the advice of the thesis supervisors.

The first token related question during the recorded interview is question three "What challenges or problems does the hardware token have to solve?". The goal of this question was to facilitate open ended thinking about the required steps for a successful offline key exchange. The required steps have been described in the section above. Regardless of the participants figuring out the required steps, each of the then following questions was focused on one specific step of the key exchange. The fourth question "How could the key be transferred onto the device?" had the goal of having the participants find ways a user can interact with the key distribution station and the token, to have the key transferred. The fifth question of "How would you transport such a token?" was aimed at different means of transportation, e.g. is there a difference between individual means of transport vs. the use of public transportation. The sixth question was the direct opposite to the fourth, namely "How can the key be used on a PC (transmitted to the PC)?". While the question may seem similar, the distinction here is, that in the case of the key being transmitted to the computer, more transmission methods are easily available.

The seventh question "What could such a hardware token look like?" was aimed towards combining the already mentioned ideas, as the participants could think of possible combinations that could work better than others. Lastly, the participants were asked to evaluate the discussed ideas according to the Now-How-Wow method, a simple way of evaluating the potential of ideas (described in section 3.2.5 Data Analysis).

Question Type	Question
Opening	1. Please introduce yourself and tell us if you have any prior experience with any kind of hardware or security tokens.
Introduction	Short description of the use case and the importance for key exchange schemes in cryptography/security
Transition	2. What challenges or problem does the hardware token have to solve?
Key	3. How could the key be transferred onto the device?
	4. How would you transport such a hardware token?
	5. How could the key be used on the PC (i.e. transferred onto the PC)?
	6. What could such a hardware token look like?
	7. Please evaluate your ideas according to the Now-How-Wow Method.
Ending	8. Do you have any remarks, suggestions, additions?

Table 3.1: Focus group question guide.

3.2.4 Procedure

The focus group was conducted online, at a time convenient for the largest amount of participants. Before the time and date was determined, the participants were asked to answer a small set of demographic questions, regarding their age, (self-identified) gender, level of education and their experience in the field of IT. Additionally, the participants had to give their informed consent on the collection and use of their data for the focus group.

The focus group itself took 90 minutes, where an audio recording of the discussion of question 2 - 9 was created. The focus group was moderated by the author of this thesis, with the goal of asking clarifying questions, answering questions regarding the question guide and keeping the discussion on topic.

After a short introduction and question 1, the participants where again notified of the required data processing and the beginning of the audio recording of the following questions. While all participants had already given their informed consent in writing, they were given the opportunity to leave, should anyone have changed their mind, prior to the start of the recording.

After the start of the recording, the moderator presented the use case and lead the discussion through the question guide, with each question having been assigned a time frame of 5 to 15 minutes.

The official part of the focus group ended with the stop of the recording and an additional opportunity for the participants to provide remarks or suggestions ”off-the-record”.

3.2.5 Data Analysis

The demographic data was analysed using *LibreOffice Calc* to calculate averages and standard deviations. A verbatim transcript of the recorded audio was created with the use of *VLC media player*, *Express Scribe*, *LibreOffice Writer* and *Microsoft Word.*

The analysis of the resulting data was two-fold. For the data from the audio recording a thematic analysis was conducted, identifying recurring themes, codifying them and categorizing them into broader categories.

The data from the collaboration tool was sorted and categorized according to the participants evaluation of the individual ideas in the Now-How-Wow method.

In that method, ideas are sorted in a 2x2 grid, with one axis representing the originality of ideas, and the other the feasibility. Ideas with a high degree of originality and high feasibility are categorized as WOW, ideas with high feasibility and low originality as NOW, as those ideas most likely already exist. In contrast are ideas with a high degree of originality, but a low degree of feasibility, which are categorized as HOW, as it may not be possible at the current moment to implement such ideas. Ideas with a low degree of feasibility as well as originality do not have an official name, but are called NO in our setup.

3.2.6 Methodical limitations

There are some obvious methodical limitations with this focus group. The first and most important one is, that due to time constraints only one focus group was conducted. This leads to only a small coverage of the subject and potential information gain, as many useful and additional inputs were probably missed. This also means, no pilot focus group was used to refine the question guide. This can lead to less useful answers from the participants, as during a pilot focus group unclear unclear questions can be identified and corrected [23]. Another limitation is based on the participants selection. While it is advantageous to have some homogeneity in a focus group, this can lead to a too narrow focus of the answers if there is not enough difference between the participants background. This can then again lead to the missing of potentially valuable input.

3.2.7 Results: Participants demographics

The final group of participants consisted of 4 persons, of which 3 identified as male and one as female. The average age of the participants was 30 (SD: 2.58), with half of them having completed a masters degree and the other half a bachelors degree. Additionally, the participants were asked to rate themselves on a scale from 1 to 5 to which degree they considered themselves experts in IT/computer science on a scale from 1 to 5, resulting in an average of 4 (SD: 0.81).

3.2.8 Results: Interview

On the basis of the transcript of the interviews audio recording a thematic analysis was performed. The participants inputs were coded to identify recurring themes and topics and then categorized into broader aspects of the hardware token.

The quotes that are provided in this section are translated from German, as the interview itself was held in German. The translations are done in best conscience, trying to stay as close as possible to the original wording while preserving the intent of each statement.

The identified categories are:

- system considerations
- technical considerations
- hardware aspects
- data transmission methods
- transportation methods
- appearance
- general attributes
- security
- comparisons

System considerations This category contains various contributions and ideas in regards to the design of the overall offline key exchange system. While this was not one of the specified goals of the questions, the participants regularly mentioned system design aspects. The first answer to question 4, the first key question (How can the key be transferred onto the device?) was to use a ”central management, that [...] transfers the key [...], manages it and distributes the device to the customer” (P_1). Following that line of thought, a few examples on how the participants themselves had experienced such a system from the point of view of an user were mentioned. This was not the first time, system design arguments were brought up, as during question 3 the participants mentioned the necessity that the issuer of the key can invalidate or reissue an already created key, if for example the hardware token is lost. During question 5 on how the device can be transported, the issue of invalidating keys was mentioned again, as well as the possibility of providing additional factors for authentication, that go further than just the possession of the hardware token and therefore the contained key.

Technical considerations: Closely related to the system considerations were the aspects based on the technical aspects of the system interfaces. Among other things, participants were concerned, that with an ubiquitous interface like USB, the device could just be arbitrarily plugged into devices with an USB port, which could pose a security risk. It was mentioned, that the use of a more restrictive type of interface, could prevent this issue, but would require additional effort e.g. in the form of additional hardware. Other technical consideration were made regarding the protection of the device against outside manipulation.

Hardware aspects: This category can be seen as having some similarities to the technical consideration, but the focus here being directly on the hardware aspects of the tokens. This includes whether the token should include displays, cameras, microphones or an audio output as well as other aspects on form or function. During the discussion the participants mentioned RFID capable chips being implanted under the skin of the user, an hardware key, that could encode the encryption key directly through its shape, just the way a normal physical key for a lock does and the possibility of using RFID capable stickers or even semi-permanent tattoos.

Data transmission methods: This is the first category mentioned here, that had a questions explicitly asking for ideas in this category, namely question 4 ("How can the key be transformed onto the device?") and 6 ("How can the key be used on the PC?"). This resulted in a wide array of answers in this category. First there are the common transmission methods, like the use of Wi-Fi, RFID. The aspects of RFID were specifically mentioned when talking about small, passive devices, like the above mentioned under-skin implanted chips. Some participants raised security concerns with those transmission methods, so an attacker can't just scan the device while passing by to extract the key.

Other transmission methods discussed were based on light or images. Examples mentioned here were QR-codes, a display flashing a specific pattern of colors or it just simply shows the key for the user to type. The use of audio signals or vibration patterns for the transmission of the key was discussed too. For audio signals one idea mentioned, the hardware token reading out the key and then the user having to type it in or the encoding of the key in an audio pattern e.g. a specific melody, which has then to be typed in by the user. A different approach was mentioned for the vibration transmission, as here the user was not supposed to type the pattern in as the PC could read the vibration pattern with a microphone. Vibration based transmission was discussed based on feasibility and originality, as it shouldn't be too difficult, as every smartphone already has an vibration motor. The possibility of the token using haptic aspects to present the key to the user, similar to braille was mentioned too. In the same vein was the idea of using memory foam and some kind of sigil for the key transmission, but that idea was not explained further.

Transportation methods Question 5 ("How would you transport such a hardware token?") was aimed explicitly towards this category. The discussion mentioned the use of courier services or postal services, as well as personal transport options. A specific instance of courier service was mentioned, when one participant described trusted courier transport with "if it is really important for the company, they won't use DHL, but send Florian from IT with a company car." (P_2). Important for the transport category, were aspects regarding the size or form factor of the device, as practicality, ease of use and inconspicuousness were mentioned.

Appearance In the discussion some aspects of appearance and transportation methods were linked closely together, as smaller and more inconspicuous devices are easier to transport, if they have to be carried by the user itself. The device should look unique, be somewhat aesthetically pleasing, not too large ("for everything larger than a smartphone, one would have to think hard if it still has a use", P_2), but also not too conspicuous ("should not be too conspicuous, to not attract too much attention", P_3).

General attributes, security These categories contain security consideration for the token or more specifically the aspect of observeability in regard to some transmission methods as some participants were concerned with a vibration pattern or audio output transmission, since a person in the same room could listen in and obtain the key. The audio output led to considerations in regard to the accessibility of the prior discussed options and the unanimous agreement on the importance of the device having a high degree of accessibility, even though this came up not until question 8, when the participants were asked to evaluate their ideas.

Comparisons The last recurring theme during the interview is comparisons, as the participants repeatedly drew comparisons to already existing technologies, or devices they already had some experience with. Among the most common comparison were comparing the hardware token with yubikeys and RSA 2FA security tokens. These comparisons were made, when talking about different kinds of user interactions with those devices, e.g ”like a Yubikey [...], plug it in, press the button, done” (P_2).

3.2.9 Results: Whiteboard collaboration

The results presented in this section are based on the evaluation of the ideas the participants wrote down for each question on the mural board. The ideas were evaluated using the Now-How-Wow method, where the participants could place stickers corresponding to the fields on the Now-How-Wow matrix on each post-it note on mural.

Evaluation	Post-It Note
Now	Smartcard + cardreader
Now	USB-drive
Now	passive device
Now	entering the ”seed” manually
Now	central management
Now	burned onto device, not modifiable, key only handed out to specific entities that can determine key via identifier
How	letting a specific order of colors flash, camera on device can read those
Now/ No	”Hardware” Key

Table 3.2: Mural Evaluation question 3 on key transmission from the key distribution station (e.g. office workstation) onto the device.

Evaluation	Post-It Note
Now	courier
Now	parcel service
How	chip under skin
How	RFID under skin
How	implanted under skin
How/ Wow	semi-permanent tattoo
Now	thrown into bag
Now	Rings/jewelry
Now	as keychain
Now	card in wallet
Wow	sticker stuck on everyday use items e.g. phone
Wow	chip in sole of shoe (compare Goldfinger[1])
Now	pants pocket
Now	suitcase

Table 3.3: Mural Evaluation question 4: "How would you transport such a hardware token?".

[1] This was a reference to the 1965 movie James Bond 007 - Goldfinger

Evaluation	Post-It Note
Now	NFC
Now	USB
Now	RFID
Now	smartcard cardreader
How/ No	Hardware Key Key in digital lock
Now	typing, if display exists
Now	IEEE 802.11ax
Now	Bluetooth
Now	Camera
Now	Device generates Number
How	Device generates sound (headphones?)
Wow/ No	Device generates Vibration
No	Device generated 3D pattern (braille code?)
How	Device generates light
How	Devices shows gestures to be repeated in front of camera
How	memory foam with pressure sensor for use with sigil

Table 3.4: Mural Evaluation question 5: "How can the key be used (transferred) onto the computer."

Evaluation	Post-It Note
Now	USB drive with buttone
Now/ How	3.5mm audio jack
Now	speakers
Now	display
Now	microphone
Now	hardware button
Now	device has to be woken actively
Wow	accessible
Now	can be carried closely at the body without bothering the user
Wow/Wow	unique
Now	rounded without [sharp] corners
Now	aesthitcally
Now	small and handy
N.A.[1]	square, useful, good ;) [2]

Table 3.5: Mural Evaluation question 6, ’How could such a token look like?’.

The following abbreviations will be used in the text: Now: N, How: H, Wow: W, No: X, and can be seen in fig.3.2 with the answers and evaluations of each question in its own table (see fig 3.2,3.3,3.4,3.5), the content of the post-it notes will be translated, as closely as possible to the original meaning, as the answers on the mural board were given in German. If more than one evaluation was given, i.e. more than one sticker was placed with a mural post-it note, both will be mentioned in the evaluation column.

The ratings on the board can only serve as an indicator for the overall rating for each post-it note, as not all participants rated all ideas on the board.

[1] no rating was given here

[2] This is the literal translation of Ritter Sports German marketing slogan ”quadratisch, praktisch, gut”

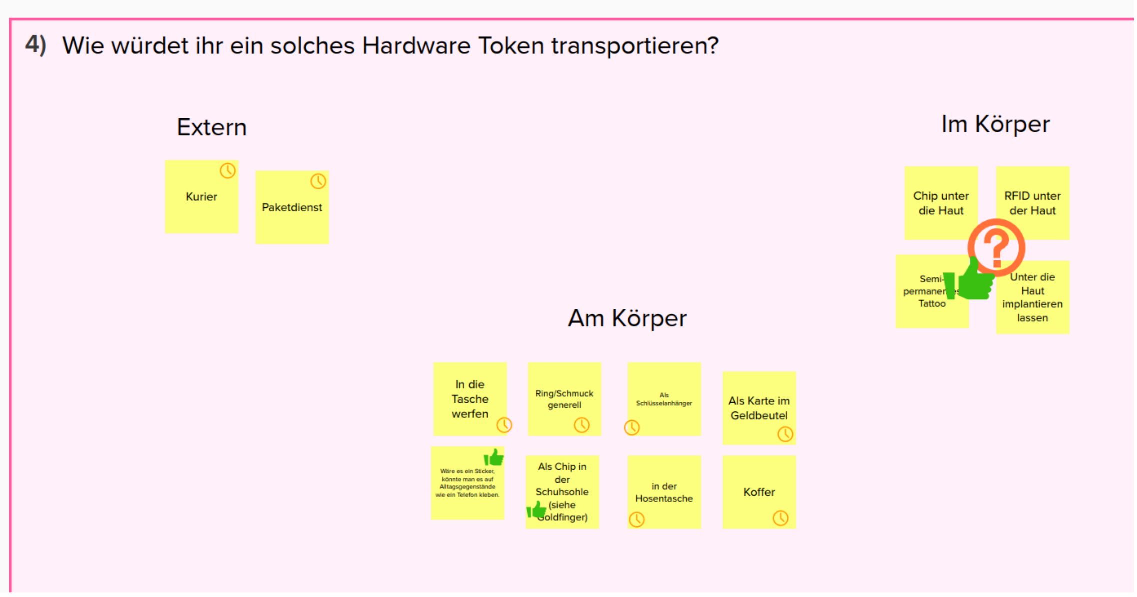

4) Wie würdet ihr ein solches Hardware Token transportieren?
Extern
Kurier
Paketdienst
Im Körper
Chip unter die Haut
RFID unter der Haut
Semi-permanentes Tattoo
Unter die Haut implantieren lassen
Am Körper
In die Tasche werfen
Ring/Schmuck generell
Als Schlüsselanhänger
Als Karte im Geldbeutel
Wäre es ein Sticker, könnte man es auf Alltagsgegenstände wie ein Telefon kleben.
Als Chip in der Schuhsohle (siehe Goldfinger)
in der Hosentasche
Koffer

Figure 3.2: The mural board for question ”How would you transport such a hardware token”, the icons show the ratings according to the Now-How-Wow method, with thumbs-up as ’Wow’, questionmark as ’How’ and the clock as ’Now’. For better readability the anwers and ratings are also shown in tab. 3.3.

3.2.10 Results: Limitations

While some methodical limitations have already been mentioned above, these can be seen directly in the results. The first one, conducting only one focus group with only 4 participants led to a narrow set of personal experiences. While the participants had different computer science backgrounds, many off them already had experience with specific 2FA hardware tokens. This may not be a problem per se, but 2 participants themselves mentioned, that their answers were influenced by their experiences. While another participant did not voice that concern himself, that same influence can be seen in some of his answers. If one were to take a close look at the transcript of the interview, that specific influence becomes even more apparent, as the moderator had to remind the participants at least 2 times, that the goal was to gather ideas for a hardware token that is used for the transport of a key, and does not generate keys itself. Even so, the discussion drifted back multiple times, to a device that would generate multiple keys.

There was not only a bias towards the RSA tokens, but the technical background of the participants led to them focusing more on the technical and implementation problems of their ideas, as well as trying to come up with possible solutions to those problems, even though this was not required by the question.

Participants mentioned that the goal of some of the questions were unclear, so they did not know on what aspects their answers should focus on. This problem could have been remediated with a pilot focus group, as mentioned in section 3.2.6.

An unclear direction of the questions, as well as the participants technical background led to ideas that were more closely linked to the key exchange system itself, e.g. when discussing central key management or the necessity of being able to invalidate keys instead of the primary goal of the usability aspects of the key exchange and the hardware token.

Lastly some participants struggled with the evaluation system, one participant having misunderstood how the axis of feasibility and originality interacted, while another participant finding that evaluation method too limiting as only two aspects were evaluated, but not if the ideas itself were even sensible solutions. This also, could have been found out and addressed with the use of a pilot focus group.

3.2.11 Discussion

Taking the methodical limitations and their direct impact on the results into account, there are still some interesting observations to make. As already mention in the section limitations, the tech background of the participants heavily influenced their answers. This can be seen with the level of detail some specific implementation aspects were discussed and the technical language that was used. While this has taken up a not insignificant time of the discussion, only very few elements of that aspect were represented with post-its on the mural board. Only one post-it (central management) was dedicated to the design of the system. The same is true for the technical considerations, which are barely represented in the collection of mural post-its. Comparing these two categories to the categories of transmission methods or hardware considerations, a stark difference can be observed just accounting for the amount of post-its for each category. An even bigger difference is seen with transportation methods, while those were discussed at length, the amount of post-its related to various transportation methods cover many examples mentioned during the discussion. Only two ideas were evaluated with 'Wow', but one of them ('chip in shoe sole, Goldfinger') could have been written down as a joke, when that example was mentioned during the discussion. The other Wow (in conjunction with How) rated idea was the use of semi-permanent tattoos. These were mentioned in the discussion as the next step after the RFID capable stickers. While no additional content on the tattoos was given, this is nevertheless a new and unique method to transport data. While in theory quite similar to the stickers or the under-skin implanted chips, one could see some advantages here, in regards to accidental loosing the device and being less invasive than sub-dermal implants. Many other hardware ideas mentioned were just already existing technologies or technologies the participants already had experience with. This could be noticed when the discussion shifted multiple times back to a hardware token that could directly generate keys (similar to existing RSA hardware 2FA tokens) together with the usability problems that occur with time based key generators. The prior experience of the participants could also be noticed in the answers to question 5, when various interesting ideas were mentioned on how the device could present the key (light patterns, images, vibration, sound) but the input methods mentioned by the participants were most of the time relying on the user to then input the key manually. Considering the length of encryption keys, the participants noted, that it could prove quite difficult for the user to correctly input audio sequences or vibration patterns, whereas the more obvious approach, of using the computers hardware, e.g. microphones or cameras to process the key directly was barely given any attention. Only when discussing transmission via vibration patterns the automated data input by the computer was mentioned in terms of accidentally introducing errors in that process by bumping into the table that the vibrating device is on.

The device generating 3D patterns was rated 'No', even though it is quite a novel approach. No further information on how this could work or be used, was given, as it was more of a side remark during the discussion.

The results on the appearance of the token deserve some attention. Many ideas related to the appearance were rated 'Now', which is not surprising as the ideas mostly reference singular already existing hardware aspects. From the discussion it became apparent, the primary concern was the size and form factor of the device, as it was mentioned multiple times by different participants, that the hardware token should be at best small enough to transport it on a keychain or directly in the pockets of one's pants. In the same vein are the examples given of having the hardware token being part of jewelry or being in the form of a sticker, that can be stuck onto the users phone. An important aspect of all the small size was that the token would be harder to loose if affixed to e.g. the users phone, as well as being potentially less attention grabbing and therefore being less likely to be stolen. Apart from that, the general consensus was that the device just has to have the appearance it needs to have based on its hardware and technical details. This may be caused by the technical background of the participants, as one could assume, technically inclined people could be more likely to accept a worse user experience, if the device still fulfills its functions properly. Even though that may be the case, the general aesthetics of the device as well as the tangibility and robustness of the device were mentioned as important, e.g. if the device has a button, it should ”feel like a button, and not like I’m destroying the device while typing” (P_2).

A nearly forgotten, but still deemed important aspect of the hardware token was accessibility. While only one example was mentioned of the ideas with the audio output being not accessible to people with hearing impairments, accessibility should still be taken into consideration for the design of the hardware token.

3.3 Selected Concepts

From the results of the focus group no strict requirements for the hardware token can be gained, apart from the mentioned aspects on size and form factor.

As size and form factor have been mentioned multiple times by various participants as being important aspects of the hardware token, these will be the primary selection criteria. An upper bound for the size has been set with the comment by P_3, to consider if anything larger than smartphone still has its use. Multiple different data transmission methods have also been mentioned as well as novel concepts (e.g. the 3D memory foam token). When choosing a token to implement for this study, the expected amount of work to implement that device have to be considered in addition to the way those tokens can be evaluated in an user study. This disqualifies ideas like the sub-dermal chip due to ethical reasons or the memory foam token due to technical reasons. An additional consideration is to choose devices that use different data transmission methods in order to increase the differences between the user interactions. If all devices use e.g. RFID to transmit the key, the user interaction would be very similar for each device and therefore probably its usability. This in turn could make the first research question, to compare the usability between device, superfluous, as the remaining differences between the devices are then only the tangible and security aspects.

The first chosen concept is that of an RFID capable sticker. The idea of a sticker on everyday items was rated ’Wow’, making it a unusual concept, with a very simple user interaction to read or write the keys onto the device. As RFID tags exist in various sizes and forms, this should be easy to implement as a prototype.

The next concept is the USB-keychain, simply taking an already existing and widely used method for data transfer and storage and see how well this device compares. USB-drives as well as keychain elements have both been mentioned during the focus group for the category of how the device could look like and for how the device could be transported. Additionally this offers a completely different user interaction and tangibility than the sticker.

The last concept probably skirts at the upper bound for size, but as devices with display and camera have been mentioned as well as the option of using QR codes to transmit data, a small device dedicated to read and display keys in the form of QR codes is chosen. This again allows for a completely different user interaction from the other two mentioned devices.

4 Implementation

In this section, a detailed overview of the implementation chosen in chapter 3 is provided. Additionally, this chapter describes two ’baseline’ devices, which represent two methods currently used for the transport of cryptographic keys or more commonly, in the case of paper, PINs and other registration codes.

For each device the user interactions to transfer a key onto the device, as well as how to transfer a key from the device to a computer, will be described.

4.1 Baseline Paper

The use of paper as a baseline is grounded in the prevalence of e.g. registration codes for online banking apps, which are frequently sent by mail and then have to be manually typed into the website of the bank. A key printed on paper could theoretically have many forms which are not just text to be typed in by the user but instead be transmitted with less user interaction, e.g. with the use of QR codes.

The description of this prototype is fairly short, as this is a regular DIN A4 sheet of paper, with a 256 bit hex encoded key printed on it, resulting in 64 characters.

While the typing of a long, random string of characters is slow and error prone, this allows this prototype to work as a verification for the user study. It can be expected for this device to perform worse if not worst of all, in the aspect of usability. If this is not the case, this means either the study design measures the wrong things or the other devices were designed or implemented with the wrong focus.

User Interaction The user interaction for this prototype is straightforward. To receive a key from the office workstation or the key distribution station, the key has to be printed out. To enter the key at the home office computer, the key has to be manually copied [by user input] into some kind of text input field.

Alternative methods would be the use of mnemonic codes to encode the key, which could be typed in easier.

4.2 Baseline Smartcard

Smartcards are already commonly used in our everyday lives, be it in the form of credit cards, national id cards or employee identification badges. The examples follow the ISO/IEC 7810 ID-1[1] format, which specifies the physical characteristics, most notably the size, of the card.

While smartcards can have different forms and sizes, if talking about smartcards in the context of this thesis, it refers specifically to smartcards with the ISO/IEC 7810 ID-1 form factor and using either a contact (ISO/IEC 7816) or a short range contactless (ISO/IEC 14443) method for communication with a card reader [33, p. 15ff, p. 297 ff]. Today many are dual-interface cards, that support both methods. This allows a smartcard to be used in a more convenient way, when less security is required e.g. building access or contactless payments of small amounts or in a more secure manner e.g. for money withdrawal at an ATM or user login for workstations in an enterprise setting.

The baseline smartcard prototype aims to emulate a contact smartcard that has to be inserted into a smartcard reader to be read or written to.

4.2.1 Technical overview

The smartcard used for the prototype in this thesis is a clone of a *MIFARE Classic 1K* proximity card, bought in bulk on amazon.[2] The main difference between the clones and authentic smartcards is that the clones allow to change the UID in block 0, sector 0, which is normally write protected and contains manufacturer data [39].

The 256 bit key (32 bytes) is written to block 8 and 9, split into two pieces of 16 bytes each, as a single write operation writes only 16 bytes at once. The key is read from those blocks using two read operations, as those return 16 bytes from one block.

[1] no primary source referenced for ISO standards, due to them not being publicly available

[2] https://www.amazon.de/gp/product/B07TVJPTM7, last access 30/08/2022

The default key for a MIFARE Classic is 6 bytes with the hex value of 0xFFFFFFFFFFFF and has not been changed. It is required for the card reader to provide the key to gain read and write access to the sectors specified in the read command.

4.2.2 User Interaction

Writing keys to the smartcard To transmit keys to the smartcard, a user needs to hold the card to the reader and wait until the transaction is finished. As the goal for this prototype is to simulate a contact card, a card reader will be used, into which the card has to be inserted first.

Reading keys from the smartcard Reading the keys from the smartcard is identical to writing the keys on the smartcard. For simulation purpose, the same card reader will be used, into which the card has to be inserted first, before the read transaction can occur.

4.3 RFID-Sticker

Small RFID tags are widely available and in use daily, often without us directly noticing them. They range from anti theft mechanisms in stores to building access tokens. Other examples include payment solutions in communal/student housing or bracelets for the lockers in public pools.

All of these examples work in a similar way, using short range electromagnetic waves to be powered and exchange data with reading device, similar to contactless smartcards, as they both follow the ISO/IEC 14443 standard for communication.

Some countries use this technology for their public transport fare cards, in which the electronics are integrated into the printed tickets. The same technology is also used in stickers (often for inventory or logistics), and therefore in this prototype. This is not a new technology and maybe from a purely technical standpoint, not even a new usage scenario, but to use these stickers as a key storage device will be a novel experience for users. Using stickers allows the users to choose the size and form of the whole key transport device, consisting of the sticker and the object the user chose to attach the sticker to. Combining these characteristics with a simple user interaction makes an interesting and usable concept for a key transport method.

4.3.1 Technical overview

The sticker used for the prototype is an ’on-metal’ sticker, using an NTAG216 chip. The model used is sold under the name ’NFC Sticker PET - On-Metal - 30 mm - NTAG216 - 924 Byte - weiß’, product number 68016, by NFC-TAG-SHOP[1]. An ’on-metal’ sticker was explicitly chosen, as during testing of a ’regular’ NTAG216 sticker, it was noticed that the sticker could not be read if directly placed on the back of a phone or on other metal surfaces. An ’on-metal’ sticker potentially allows a more flexible and creative placement of the sticker by a user.

Notable features of the sticker include the option of setting a 32 bit password (default 0xFFFFFF) to access the memory areas on the sticker as well as 888 bytes of user writeable memory.

The 32 bytes of the key are written to page 6 - 13, using eight write operations, as each write operation can only write four bytes to one page [40]. The key is read using two read operations, as each read operation returns the content of four pages.

As this tag is compliant with ’NFC Forum Type 2’ standard which in turn is based on the already mentioned ISO/IEC 14443 standard, some modern smartphones can automatically detect these tags and try to read NDEF (NFC Data exchange format) records from the tag. If the sticker is stuck on the back of such a smartphone, the phone informs the user that a new tag has been detected and which content it contains. This does not happen all the time, but is dependent on the

[1] https://www.nfc-tag-shop.de

specific sticker placement and may vary between different smartphone models. Testing this with a Pixel 6 a sticker placed lower on the back did not automatically get picked up by the automated tag detection, but placed in the middle of the back the phone registered a new tag (see fig. 4.1). In the test case of the Pixel 6 the phone had a full screen message pop up every time the phone was unlocked, which could be annoying. In the unlikely case, that there is no spot on a phone, where the sticker would not get picked up, the phone could not be used as the physical carrier object for the sticker, unless NFC function on the phone is turned off. Doing this would also turn off other NFC based features, like Google Pay. In comparison, the same behaviour was not observed when testing sticker placements on the back of an iPhone 12 mini. Not tested was the impact of a sticker placed on the back on features like contactless charging.

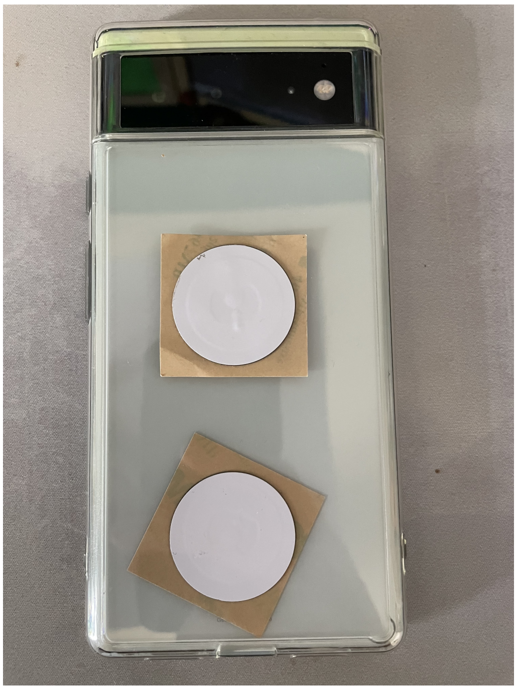


Figure 4.1: Two NTAG216 PET 30mm stickers placed on the back of a Pixel 6. The lower placed sticker is not automatically detected by the smartphone.

4.3.2 User Interaction

Writing keys to a sticker To transmit keys to the sticker, a user needs to hold the sticker close to the reader and wait until the transaction is finished.

Reading keys from a sticker Reading the keys from the RFID-Sticker is identical to writing the keys to the sticker.

4.4 USB Keychain

This prototype was chosen due to the small size of many USB drives and the often existing option of easily attaching it to a keychain, making it an ideal candidate for a small, unobtrusive keychain element, which should not stand out on most key rings.

While the original concept idea included a fingerprint scanner for the USB drive, this security feature was dropped in the final prototype, to keep all devices comparable among each other. As the other devices have no directly user interactable security features, having a fingerprint scanner on this prototype would probably skew the results on perceived security, as it can be expected that the device with visible security features will get rated better than the other devices without.

4.4.1 Technical overview

The USB memory stick used was a Intenso Micro Line 8 GB, bought via amazon[1]. This specific device was chosen because of its size. Its especially small size of 18 x 15 x 7 mm (L x W x H) [18] should allow it to not add too much bulk to a users keychain.

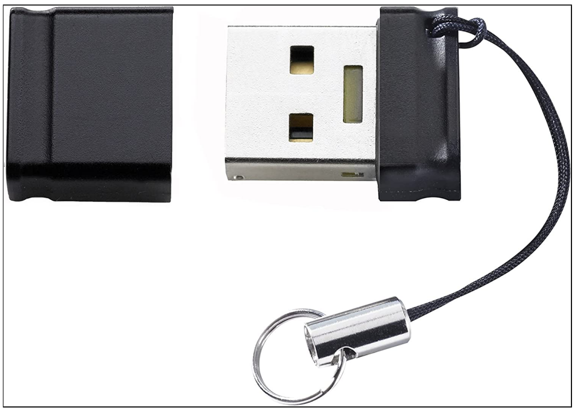


Figure 4.2: The Intenso Micro Line USB drive is smaller than many common keys or keychain elements (Image source see footnote 1).

4.4.2 User Interaction

Saving Keys on the USB drive The user interaction to save a key on the USB drive is similar to saving any other file from an application to the file system of the computer. First the user has to plug the USB drive into an USB port. Under the assumption that modern operating systems automatically mount the drive on their local file system, the USB drive can be used afterwards. In the application used for the issuing of keys, the user selects the 'save key' option. This opens a dialog for saving the file. Now the user has to navigate to the USB drive and then click the save button. Before clicking the save button the user has the option of customizing the file name, as is generally the case when saving files. After the file has been saved, the user can safely eject and unplug the USB drive.

Opening keys from the USB drive The user interaction to open a key file from the USB drive is similar to the user interaction of saving the key file in the first place. First the user has to plug the USB drive into an USB port. Now a user can use the 'Import Key' option to open a file selection dialog. Here the user navigates to the USB drive, selects the file and chooses the import option. After that the USB drive can be unplugged.

4.5 QR-Device

This section provides an overview of the key transport device 'QR-Device'.

This device was created directly from the suggestions made by the focus group participants. It combines the idea of transmitting keys via QR codes with the size constraint of 'not bigger than a smartphone'. The device itself is fairly simple, as it uses a camera to scan QR codes, decodes the key encoded in the QR code and saves it to local storage. To transfer the keys to the computer,

[1] https://amazon.de/-/en/Intenso-Micro-Memory-Stick-Black/dp/B07VC3Z4KT, last accessed 30.08.2020

a key can be selected, is then displayed as a QR code, and can then be scanned by any suitable application. The device itself has no other communication methods and is completely air-gapped.

This concept is not novel, as there are a few similar devices existing, primarily in the form of cryptocurrency hardware wallets or signing devices. Cryptocurrency wallets do have some similarities with the offline key transport devices in this paper, namely that they have the same goal, keeping keys secure. Three different cryptocurrency devices use a system, similar to the QR Device. First is the keystone wallet [1], a ready to use, commercially available hardware wallet using QR codes to sign and verify transactions. The other two devices are the Specter DIY hardware wallet [2] and the transaction signing device seedsigner [3].

4.5.1 Implementation

The first idea for the implementation of the QR-Device prototype, was to use one of the above (sec. 4.5 mentioned cryptocurrency hardware wallets with a custom application that fits the user interactions of the simple home office use case proposed in chap. 3.1.

This was not possible for the keystone wallet, as it does not support the option of adding custom firmware [4]. The other alternative, building and modifying as Specter DIY hardware wallet, was also not possible, due to the unavailability of some of the required hardware components, specifically the *STM32F469I-DISCO* hardware board.

As a workaround, it was decided to simulate the QR-Device using an android application on a generic android smartphone. The phone chosen for the QR-Device is a Samsung Galaxy S7, running Android 8, with turned on Airplane Mode, ”Do-not-disturb” and all notifications disabled. The app itself is set up to run as pinned app, meaning during that time the home button and switching between open apps is disabled.

The app itself was developed using Android Studio Chipmunk 2021.2.1 Patch 1, targeting SDK version 26, minimum SDK version 24. The only non standard google library used was ’com.journeyapps:zxing-android-embedded:3.6.0’[5], a library to decode QR codes. The app required the following permissions on the device: ’android.permission.CAMERA’, ’android.permission.READ_EXTERNAL_STORAGE’, ’android.permission.WRITE_EXTERNAL_STORAGE’ and ’android.permission.MANAGE_-EXTERNAL_STORAGE’. These are required for the app to access the camera to scan QR codes, to save the keys to internal storage as well as to save timestamps of the user interactions to external storage. Additionally each activity explicitly hides all status icons in the android status bar.

The app consists of the following five activities:

1. ParticipantID
2. MainActivity
3. AddKeyActivity
4. InteractionSuccess
5. ShowKeyActivity

ParticipantID This activity is used to set a new participant ID for logging timestamps.

[1] https://keyst.one/

[2] https://github.com/cryptoadvance/specter-diy

[3] https://seedsigner.com/

[4] This was clarified by Lixin Liu (Twitter: @BitcoinLixin), CEO of Keystone, through personal contact via discord

[5] https://github.com/journeyapps/zxing-android-embedded

MainActivity This is the apps home screen and provides a list of all available keys saved on the device, as well as the option to add new keys (Fig 4.3a).

AddKeyActivity This activity is used to scan a new key and save it to internal storage (fig. 4.3b). To add a new key, the user has to position the QR-Device in front of the QR-code key, with the camera pointing towards the key. The camera preview in this activity supports the user in positioning the device correctly.

InteractionSuccess This activity is displayed, after a key was successfully read by the app. On clicking 'DONE' the user is returned to the MainActivity (fig. 4.3c).

ShowKeyActivity This activity displays the QR code of a key, so that the user can present the key the receiving device (fig. 4.3d).

4.5.2 User Interaction QR-Device

Reading Keys with the QR Device To read a key with the QR-Device, the user first has to tap the 'ADD NEW KEY' button on the apps main screen. This opens the AddKeyActivity, where the user now has to point the camera of the QR-Device towards the QR-code containing key the user wants to scan. If the device is positioned correctly, the key is automatically read and saved to internal storage. After the key is saved, a success message is displayed.

Reading keys from the QR-Device To transmit a key to an application, the user first has to select the key, that is going to be transmitted, from the key list. This opens the ShowKeyActivity, where the key is shown as a QR code on the display of the QR-Device. The user now has to position the QR-Device in front of the a receiving devices camera, with the screen facing towards the camera. This then allows the receiving device to scan the QR code, and decode the key contained therein.

5 User Study

After suitable prototypes and the baselines have been chosen, a user study was conducted to compare the prototypes regarding aspects of usability, tangibility and security.

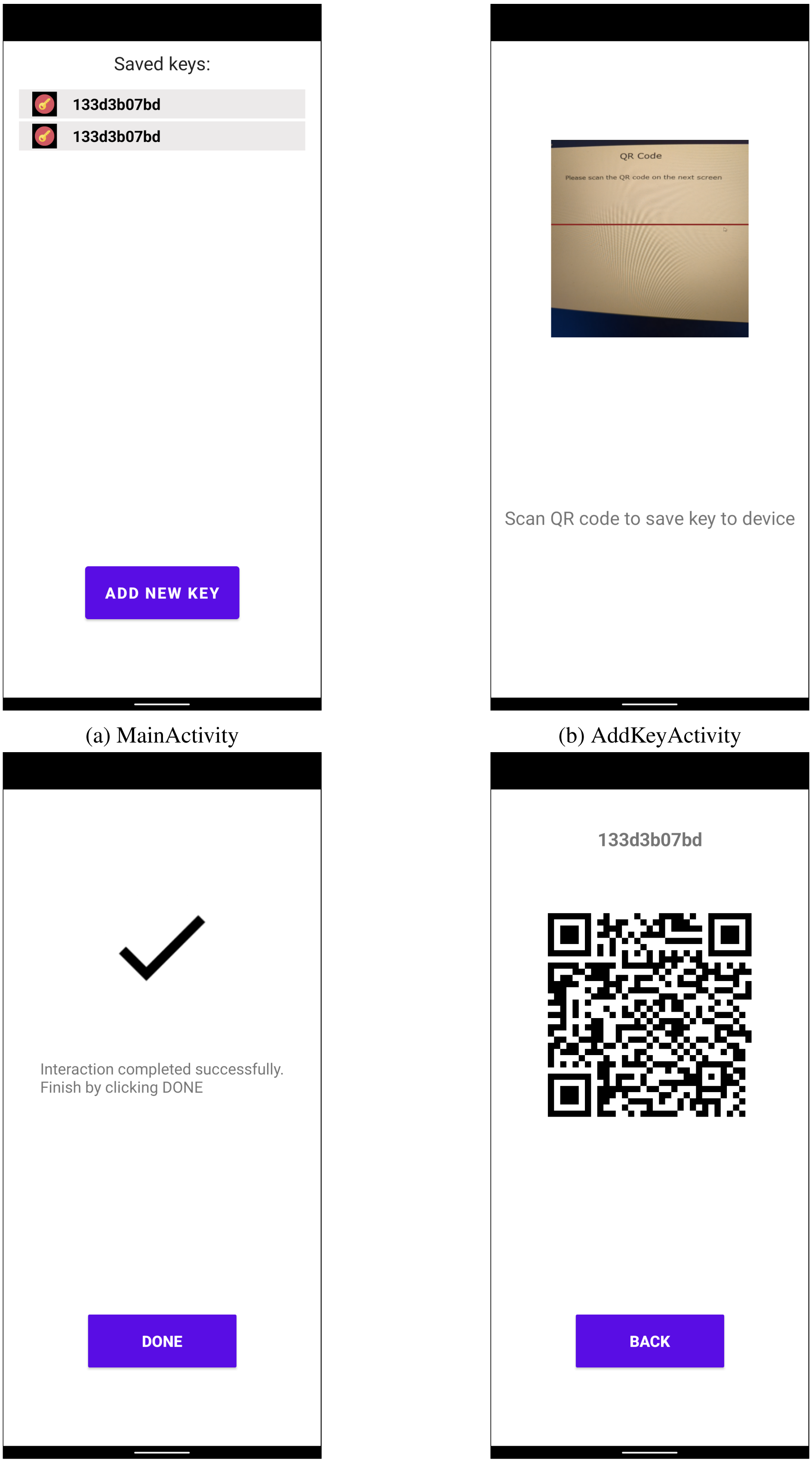
5.1 Study Design

The study was designed as a repeated measure within subjects study. Here, each participant has to use each prototype to fulfill the same tasks. To avoid negative ordering effects, the order of prototypes to test was determined using a balanced latin-square matrix with the last digit of the participant id determining the order of evaluations.

For each prototype, the participants have to solve two tasks. First, using the prototype to save a key on the prototype and second after transporting the prototype, transfer the key from the prototype back to a computer.

5.2 Apparatus

This section describes the experimental setup of the study, as well as the supporting software and hardware build to conduct the study and allow the participants to interact with the prototypes.



Saved keys:

133d3b07bd

133d3b07bd

Scan QR code to save key to device

QR Code

Please scan the QR code on the next screen

Interaction completed successfully.
Finish by clicking DONE

133d3b07bd

(a) MainActivity

(b) AddKeyActivity

(c) InteractionSuccess

(d) ShowKey

Figure 4.3: All activities of the QR-Device app.

5.2.1 Experimental setup

The user study was conducted in a temporarily vacant office on the floor of the Usable Security and Privacy Group in the building of the Research Institute CODE.

On two desks, two workstations facing each other, have been set up. The workstation on the right, simulating the office (in comparison to the home office on the left) was a Microsoft Surface (exact model unknown), with the card reader connected through a USB-hub power supply combination. This left the USB port on the right side of the surface free to use for the USB-Keychain prototype. The smartcard reader was placed left to the screen on the side closer to the camera.

On the opposite desk, the home office workstation was set up, using a Lenovo Thinkpad 13 (2nd Gen.), with the smartcard reader on the right side of the laptop, closer to the camera. The placement of the smartcard readers was chosen to allow for an easier analysis of the interaction with the card reader with the recorded videos.

The camera used for recording was a GoPro Hero 3+ Black Edition, recording in 1080p, 30 fps, wide angle mode, mounted on a hatstand at a height of ca 160-170 cm.

A third computer was setup, outside of the camera view, for the sole purpose of answering the questionnaires. The setup is outlined in fig 5.1.

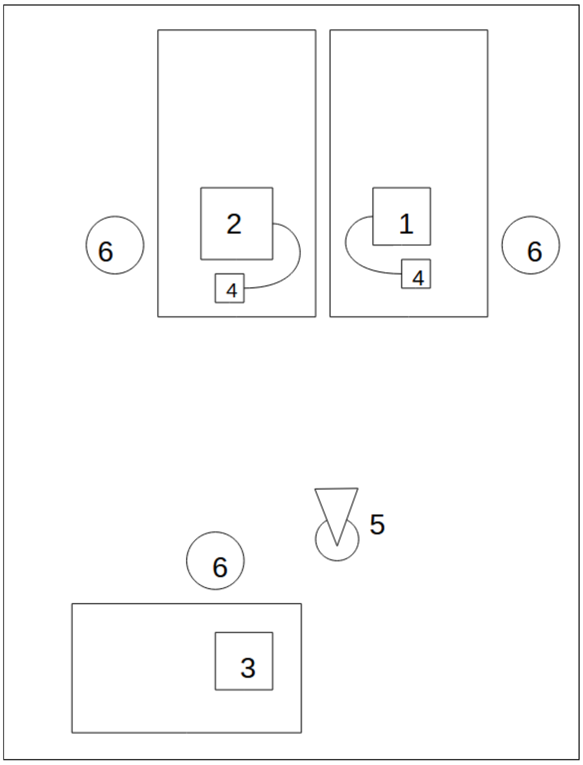


Figure 5.1: Setup for the user study. (1) Office PC, (2) Homeoffice PC, (3) Questionnaire PC, (4) Card reader, (5) Camera, (6) Desk chairs (not to scale).

5.2.2 OfflineKDApp

In this section the OfflineKDApp is described. This app was specifically created for the user study to provide the participants with a way to start the interaction tasks for each prototype, have a short description of the current task and receive feedback when an interaction was completed.

Implementation The app was created using python 3.7. It consists of a simple GUI, using the 'tkinter' library. The GUI consists of multiple tkinter frames, stacked on top of each other. If a new frame is needed, the app raises the specific frame to the top of the stack. Each prototype has one frame for the transmission of the key onto the prototype as well as for the transmission from the prototype to the computer. Additionally there is a frame to show when a task was completed. Depending on if this was during the first run (test) or second (measure), a different message is displayed. On the home office computer, after the first run, the participant is informed the test run is completed and asked to move back to the office computer where the second run can be started. After the second run, the participant is informed that this task is completed and asked to answer the questionnaire for the prototype they just used. On the office computer, the message is nearly identical for both runs, informing the participant of the completion of the run and asking the participant to move on to the homeoffice computer.

For the frames that are used to start the read or write operations for the smartcard and RFID-Sticker, a small protocol handler was implemented. During the startup of the app, a connection to the smartcard reader is established and if a read or write request is triggered through user interaction a thread is spawned to communicate with the smartcard reader. The results of the communication thread are handed back to the main thread via queues, which are periodically polled by the main UI thread. If a "success" status message is found in the queue, the GUI updates accordingly with a interaction success message.

5.2.3 Card Reader

To allow the reading and writing of keys for the smartcard as well as for the RFID-Sticker a reading device was needed. For that an arduino uno r3 (specifically an elegoo uni r3) was used, in conjunction with a generic Arduino MFRC522 RFID reader (see fig. 5.2). This type of reader is available in many arduino starter kits and can be found online for around 3€. This RFID module is compatible with both (older) MIFARE cards, as well as NTAG21x cards, so it can be used to create a single smartcard reader that can be read both the RFID-Sticker and the smartcard [41].

To emulate a real smartcard reader, a case for the reader module was designed using tinkercad[1], so that the smartcard can be be inserted for the read/write process (see fig 5.3).

The design is based on a a RFID module case design by Sharklade [2], licensed under CC BY-NC-ND 4.0. The final 3D printed and assembeled case is depicted in fig. 5.4.

Implementation The card reader was implemented using the Arduino IDE 1.8.19. The library used for the MFRC522 module is the MFRC522 library, version 1.3[3]. While not all features of NTA21x tags are supported, simple read and write operations are.

For the communication of the arduino with the OfflineKDApp a simple text based protocol over a serial connection was designed.

The desktop app can request either a read or a write operation, in the following format: "(R|W)(s|r)([0-9a-f]{64})?". The first character specifies if a read (R) or write (W) operation is requested, the second if its for the smartcard (s) or the RFID-Sticker (r). If a write operation is requested, the characters following the target identifier (s or r) is a hex encoded 256 bit key, that is to be written to the card or sticker. After such a message is received, the card reader replies with either a 'READ_ACK' or 'WRITE_ACK', to communicate to the application that it is waiting for a card or sticker. If a card or sticker is inserted, various debug messages are sent to the app and logged. After a read/write operation for the card or sticker has been finished, the card reader

[1] https://www.tinkercad.com/

[2] https://www.thingiverse.com/thing:4093415

[3] https://github.com/miguelbalboa/rfid

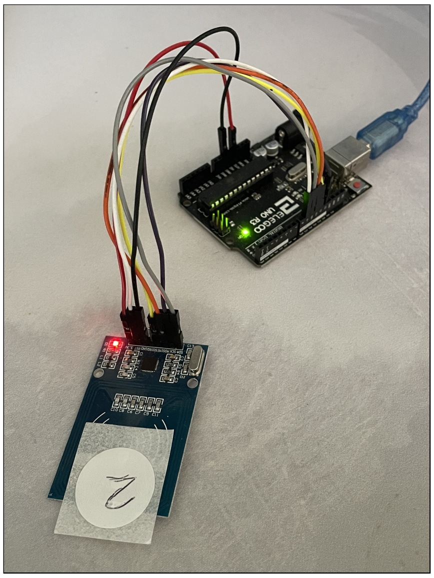


Figure 5.2: Arduino RFID module to be used as card reader, connected to arduino with an RFID sticker prototype placed on the reader.

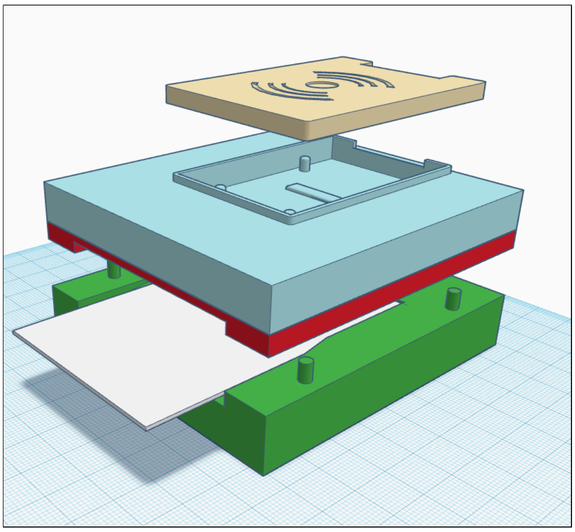


Figure 5.3: Final design of the .stl files for the individual parts of the case, created with tinkercad.

send either "QKD:([0-9a-f]{64})" with the read key for the read request, or ’DONE’ for the write request.

Various errors can occur, if the card or a sticker could not be read successfully. This can be caused for example if the sticker is barely held inside the contactless communication range, or is removed too fast before the operations are completed. In these cases, the card reader sends an error message to the app and triggers a full reset of the arduino. The app waits for three seconds, before resending its last request. While this is not the cleanest solution for handling errors, during testing this was the only method that would reliably complete the requested read or write if the card was simply inserted again, or the sticker removed and placed again on the reader.

5.3 Participants

The participants for this study were recruited through the study participation mailing list of the LMU Mediainformatics department, the LMU Munichs infoservice, social media posts and direct

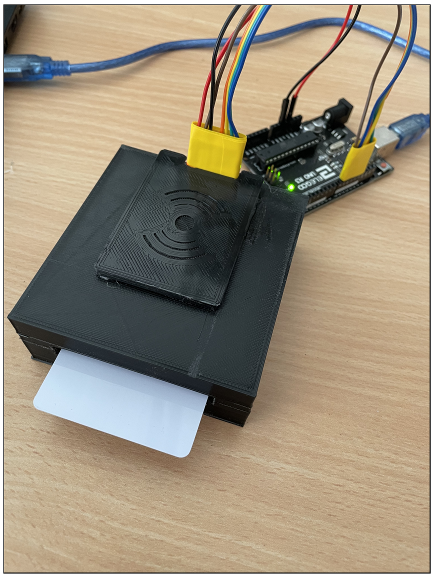


Figure 5.4: Final 3D printed smartcard reader case with inserted smartcard.

contacts. To qualify for this study, the participants had to have worked in a home office situation before and have some English skills, as the questionnaires have not been translated into German due to missing official or validated translations. A compensation of either 10€ cash or 1 MMI study credit was offered for the participation.

5.4 Question Guide

This section provides an overview of the different questions selected for the various parts of the user study.

5.4.1 Demographics and Security Attitude

The standard demographic data we are collecting in this section are age, gender, highest level of education and employment status. Additionally, due to the home office use case, participants are asked how often they currently work from home. We also ask if the participants job description would fit the description of an IT-administrator or IT-support technician. This question is asked, as it could be interesting to see if prototypes are evaluated differently, when viewed not from a pure user perspective, but from the point-of-view of an IT-administrator that has to administrate and answer user support requests for these devices.

To assess the general technology affinity of the participants, each participant answered the ATI questionnaire [16].

Next we included two questionnaires on security attitude and behavior, namely SA-6 [15] and the Security Behaviour Intention Scale (SeBIS) [14], to gauge the general level of security mindedness.

As we not only want to measure the perceived security of each prototype, but also the level of trust placed in each one, two questionnaires to measure faith in general technology and the trusting stance on general technology. Both were taken from [26] and rated using a 5 point likert scale.

Lastly the participants answered the Need for Touch scale [32], as physical interaction and the tangibility of the prototypes is one of the three main aspects of this thesis.

5.4.2 Device specific questionaires

A larger set of existing questionnaires and and self developed questions was compiled into a device specific questionnaire, that has to be answered for each prototype. Overall, the device specific part of the questionnaire contained 39 statements/questions per prototype.

Usability To gain a rough estimate of the usability of each prototype, we used the System usability scale (SUS) [9] questionnaire in conjunction with the raw NASA Task Load Index (RTLX) questionnaire [21]. The SUS questionnaire contains 10 items, scored on a 5 point likert scale. For the final score calculation, each item can count for up to 10 points, 2.5 points for each rating above 1 (e.g. rating of 1 = 0 points, rating 2 = 2.5 points, etc ...). Negatively worded items (2,4,5,6,10) are inverted for the counting of the score. RTLX was used instead of the weighted TLX to reduce the amount of statements the participants had to evaluate for each statement, as the 15 pairwise comparison of task load factors for each device was left out. The rating of six items e.g. mental load, physical load etc., was done on a 21 point likert scale with each point above the first counting for a score of 5 for that item. For the final score the average over all item scores is calculated. The score of item 4 (Task success) was inverted, as in the original TLX questionnaire a lower value means a lower (better) score. This could not be sufficiently be modeled in the used survey tool, so in our case a higher rating on the likert scale meant a better overall score for that item.

Tangibility To gain a measurement of the effects tangibility a small set of questions was devised. The questions were rated on a 5 point likert scale, and evaluated on answers counts per item.

The following questions are gauge the effects of tangibility:

1. I am aware of the consequences of losing the prototype.
2. Being able to hold the prototype in my hands, gives me a feeling of being in control
3. I would prefer this prototype to a purely digital alternative (e.g. key exchange over the internet)

Form factor Closely related to tangibility is the aspect of form factor of the devices. The questions are relatively broad to cover multiple wider aspects of the devices.

1. I believe this prototype could get easily lost
2. I believe this prototype is easily transportable
3. I believe this prototype is handy
4. I believe using this prototype is fun

The first two question are aimed towards the purely physical side of the prototypes, while the later two questions are aimed towards the users feelings when using the prototypes.

Security The last set of custom questions is aimed towards the perceived security of each prototype. It is important to note here that the participants do not get any additional information on the security of the device. The exception is the QR-Device, because even though regular smartphone interaction elements (home button, volume button etc.) are hidden, it still bears a close resemblance to a smartphone. So the information provided is that the QR-Device explicitly is a dedicated air-gapped hardware device, without any means of communication outside of reading and displaying QR codes. This means that for answering the following questions, the participants have to rely on prior knowledge of potential security features of the devices and their gut feeling on how the device probably works and how secure that mechanism is.

The statements to rate are as follows:

1. I am aware of the consequences if another person gains access to the prototype.
2. I believe the prototype is keeping the keys secure
3. I feel the transmission of the keys to and from the prototype is secure.

Human Computer Trust Scale To measure trust in the individual prototypes we used the HCTS questionnaire. This questionnaire directly allows for customization for specific use cases [19]. The final statements are as follows:

1. I believe that there could be negative consequences when using the system.
2. I feel I must be cautious when using the system.
3. It is risky to interact with the system.
4. I believe that the system will act in my best interest.
5. I believe that the system will do its best to help me if I need help.
6. I believe that the system is interested in understanding my needs and preferences.
7. I think that the system is competent and effective in transporting cryptographic keys.
8. I think that the system performs its role as a cryptographic key transport device very well.
9. I believe that the system has all the functionalities I would expect from cryptographic key transportation device.
10. If I use the system, I think i would be able to depend on it completely.
11. I can always rely on the system for the transport of cryptographic keys.
12. I can trust the information presented to me by the system.

For the evaluation, we use the same method as used in [35]. The statements are rated on a 5 point likert scale. For the evaluation the sum of all ratings is calculated, with the negatively worded items 1-3 inverted. This results in a score between 12 (very low) and 60 (very high). In addition to the overall score, the average is computed over the subscales 'perceived risk' (item 1-3), 'benevolence' (item 4-6), 'competence' (item 7-9) and 'trust' (item 10-12).

5.4.3 Interview Questions

For the exit interview, we devised a set of questions to gain additional insights on usability aspects in the context of everyday use and security perceptions for each prototype. The interview was conducted in German, so both the original German questions as well as an English translation will be presented here.

1. Würde es dich stören, diese Devices im Alltag zu benutzen? Wenn ja, warum? Wenn nein, warum nicht?
2. Würde es dich stören, diese Devices zu transportieren? Wenn ja, warum? Wenn nein, warum nicht?

3. Wie könnte die Benutzerfreundlichkeit der Devices verbessert werden?
4. Gab es ein Device, welches du als besonders sicher empfunden hast? Wenn ja, welches und warum?
5. Gab es ein Device, welches du als besonders unsicher empfunden hast? Wenn ja, welches und warum?
6. Wie könnte die Sicherheit der Devices verbessert werden?
7. Würde es für dich einen Unterschied machen, wenn die Devices zusätzliche Sicherheitsmerkmale hätten, wie z.B. einen Schutz durch Fingerabdruck oder PIN?

English translation:

1. Would it bother you to use these devices in your day to day life? If yes, why? If not, why not?
2. Would it bother you to transport these devices? If yes, why? If not, why not?
3. How could the usability of these devices be improved?
4. Was there a device you thought was especially secure? If yes, which one and why?
5. Was there a device you thought was especially insecure? If yes, which one and why?
6. How could the security of these devices be improved?
7. Would it make a difference for you, if these devices had additional security features, like fingerprint or PIN protection?

5.5 Procedure

Before the start of the study, the supervisor entered the participant ID in the QkdTestApp running on both workstations, as well as in the questionnaire on the questionnaire PC and the QR-Device. After a short introduction on the general content of the study, the participants were reminded of the recording of the study and the option to leave anytime without negative consequences. In the next step, the participants filled out the questionnaire on demographics and the other user specific questions. After that the recording was started and the home office use case was presented to the participant. After making sure the scenario was understood, the experimental setup was explained with the following information.

For every prototype, there are two runs. In each run, two tasks have to be performed. The first task is the transmission of a key from the office workstation to the prototype. After moving to the home office computer with the prototype, the second task is the transmission of the key onto the computer. Each task begins with a click on the start button, performed by the participant. The participant can then interact with the prototype, e.g. inserting the smartcard, scanning the QR code etc. After that interaction is done and a key has been transmitted a interaction complete message is displayed. Now the participant can finish their interaction with the prototype, e.g. removing the smartcard from the reader or unplugging the USB drive (interaction during task 1 depicted in fig. 5.5). After the interaction with the prototype is finished, the participant clicks the 'DONE' button, which concludes the task.

As described in 5.2.2 a task completed screen is displayed, informing the participant of the next step. In addition to the information displayed, the supervisor also actively informs the participant of the next step.

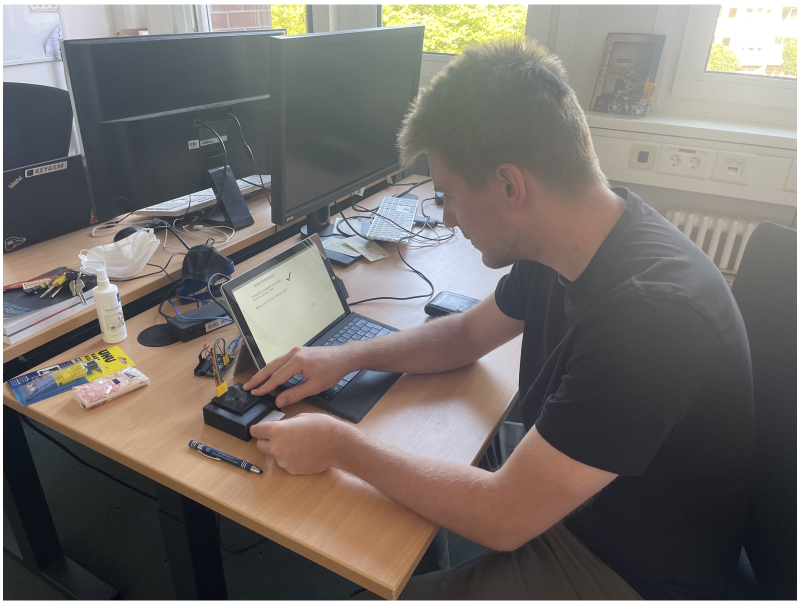


Figure 5.5: A participant interacting with the smartcard prototype and card reader at the office workstation to transfer a key to the smartcard during the first task of the study.

The first run is a test run, to allow the participant to get to know the prototype, the interactions with it and the desktop app or card reader. To start the first run, the participant takes a seat at the office workstation, where the screen for task 1 for the prototype is displayed. The participant then gets a short explanation on the function of the prototype. For the prototype Paper baseline, during task 1, no action is required by the participant, as the sheet of paper with the printed key is handed directly to the participant. For the prototype RFID-Sticker, the participant is now asked to select one object, they would like to place the sticker on. They are then handed the RFID-Sticker, combine with two stripes of electrical tape, to stick the sticker on their chosen object. In the case of the USB-keychain prototype, the participant is asked to attach the USB-keychain to their personal keychain. After that, the participant can start the first task with a click on the start button. After finishing both tasks the first run is completed.

For the second run, the participant moves back to the office workstation, clicks next on the task completed screen and can start the first task again. After repeating the sequence from the first run, both tasks are completed and the second run is done.

During the second run, the supervisor selects the currently tested prototype on the questionnaire screen, so the participant can directly begin with the evaluation after finishing the second run.

The participant now fills out the device specific questionnaire for the current prototype. Meanwhile the supervisor selects the next prototype to interact with on both workstations.

When the questionnaire for that device has been completed, the participant again takes a seat at the office computer, and the whole process begins anew with the next prototype.

The last step, after all prototypes have been tested and their corresponding questionnaires were answered, is a short interview on the usability and security of the prototypes as well as possible improvements.

6 Results

In this section an overview of the results of the user study is provided. First the tools and methods to analyse the resulting data is described, followed by the data gathered on the participants. The rest of the results is divided into quantitative and qualitative analysis of the collected data. In the former category the questionnaire results are analysed, while in the later section the participants interviews as well as additional data gathered from the video recordings are described.

6.1 Data analysis

The video files recorded by the *GoPro Hero 3+ Black* were combined into single video files, containing the whole study run for each participant, using *Microsoft Windows Photos*. This was necessary, as the GoPro automatically splits recordings into chapters of around 4gb of size. Additionally, during some participants runs (e.g. P_8, P_{11}, P_{15}), the camera randomly stopped recording and had to be started again by the supervisor. This has led to a few seconds of video and audio loss for that participant. The impact on the reliability of the videos can be considered low, as most times this happened when the participant was not in front of the camera, therefore not interacting with the prototypes or answering the interview questions. For P_{12} a few seconds of the interview were lost, due to camera issues, this also led to the 6th interview question having been skipped for that participant. The video recordings were only for qualitative analysis and will not be publicated.

Using *VLC Media Player*, mp3 audio files were extracted from the videos, to be used with the free (non-commercial) version of *Express Scribe* to transcribe the interviews. Before the transcription of the interview only the last few minutes containing the interview were extracted from the audio files using *Audacity*. The interviews were then transcribed using *Express Scribe* and coded using *Google Sheets*.

The questionnaires and the timestamp logs were analyzed using the python library pandas and jupyter. The tables and figures were also generated using the built-in methods of pandas.

6.2 Participants Demographics

A total of 16 participants were recruited for the user study, 9 identified as male, 7 as female. The mean age was 27.38 (SD: 2.6, median: 27), the age distribution is depicted in fig. 6.1. All participants had an university degree, 8 had a bachelors degree and 8 a masters degree (or equivalent).

The current occupation of the participants is depicted in tab. 6.1.

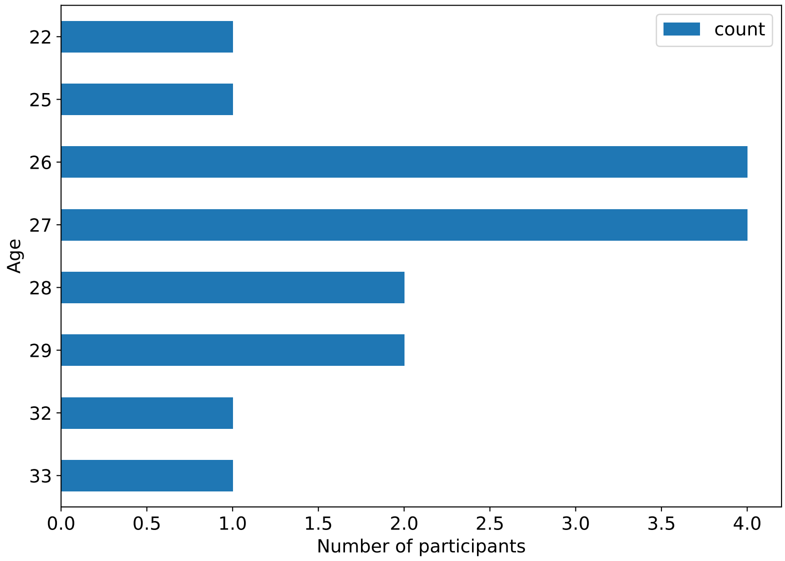


Figure 6.1: Age distribution of N = 16 participants.

The frequency of working in a home office context is depicted in fig. 6.2. The agreement with the statement of ”My job largely consists in tasks typically executed by an IT administrator (e.g.,

Occupation	count
Employee	10
University student	5
Civil servant	1

Table 6.1: Occupation status of participants.

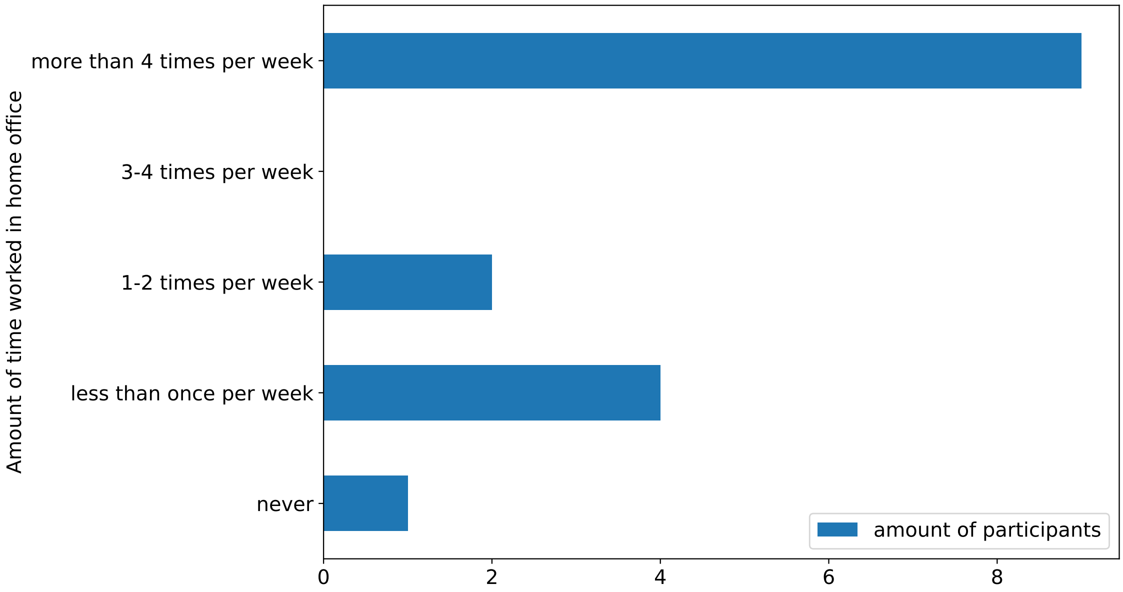


Figure 6.2: Self reported amount of time per week spent in home office.

administration of co-workers' soft- or hardware)." is depicted in fig 6.3. As most participants jobs did not consist of IT administrative work, the evaluations of the prototypes is primarily done from a user perspective, without the potential cost of administrating such systems in mind.

The mean scores of ATI, SA-6, SeBIS, Faith in general technology and trusting stance in technology are displayed in tab. 6.2.

Questionnaire	mean	SD	median
ATI	3.36	0.9	3.72
SeBIS	3.41	0.32	3.43
SA-6	3.18	0.60	3.08
Faith gen. Techn.	3.31	0.6	3.5
Trusting Stance gen. Techn.	3.23	0.68	3.33

Table 6.2: Technology affinity, security attitudes and trust in technology.

The mean scores for the Need for Touch scale and the two sub scales instrumental and autotelic are visible in tab. 6.3. The instrumental sub scale describes the "goal driven utilitarian form of instrumental touch" e.g. touching with the intention of gaining information about the object, whereas the autotelic scale is more enjoyment focused [32].

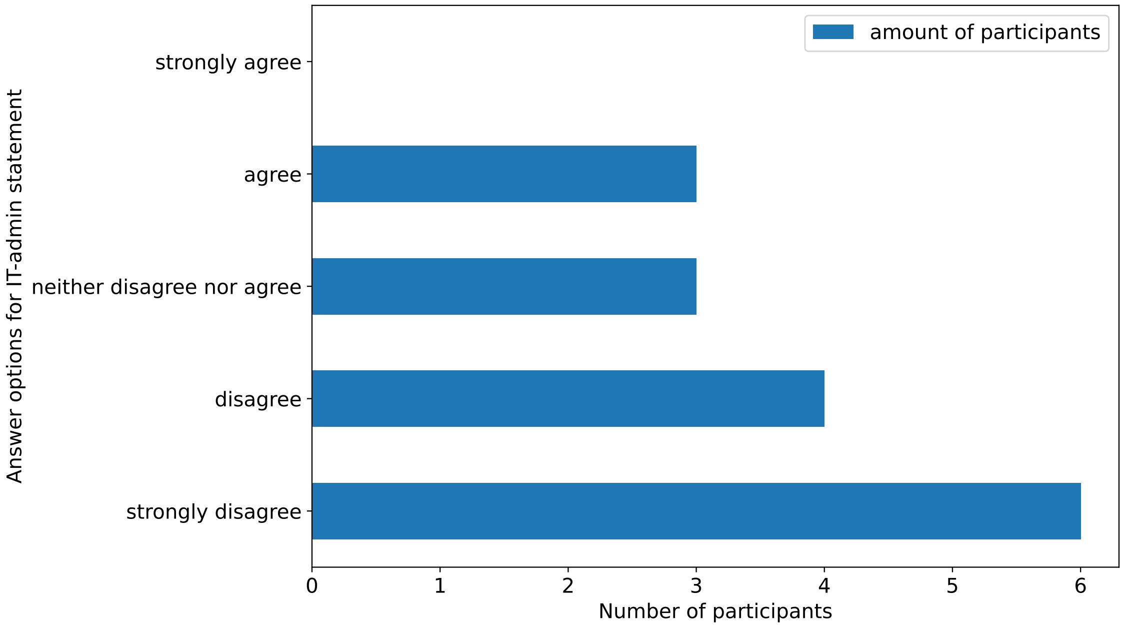


Figure 6.3: Participants agreement with the statement their job consists of mainly IT administrative or IT support work.

Scale	mean	SD	median	items counted
Full NfT	-0.38	18.24	2.5	12
Instrumental subscale	0.81	10.04	3	6
Autotelic subscale	-1.19	10.38	-1	6

Table 6.3: Need for touch and affiliated subscales. Ratings can go from -36 to 36 (full scale) and -18 to 18 (subscales).

6.3 Quantitative Analysis

This section presents the results that can be extracted from the questionnaire for the prototypes and the timing data collected during the study.

6.3.1 Usability

This section contains an overview of all quantitative data gathered that can be used to assess the usability of the prototypes.

SUS The primary measurement for the usability of the prototypes is the result of the SUS questionnaire, higher scores implies better usability. The resulting SUS score for each prototype are depicted in fig.6.4. The mean score with SD and median is shown in tab. 6.4. According to the adjective ratings for each score developed in [5], the prototypes have the following SUS ratings: Paper: ’OK’, Smartcard: ’Excellent’, RFID-Sticker: ’Excellent’, USB-keychain: ’OK’, QR-Device: ’Good’.

RTLX An other measurement that can be used as an indicator of usability is the task load index, e.g. how difficult is it to perform a specific task with that system. A boxplot of the raw task load index is depicted in fig. 6.5. Notable here is the outlier for the RFID-Sticker with 43.33, where as the mean is 5.31 (SD: 11.38, median 1.25), as shown in tab. 6.5.

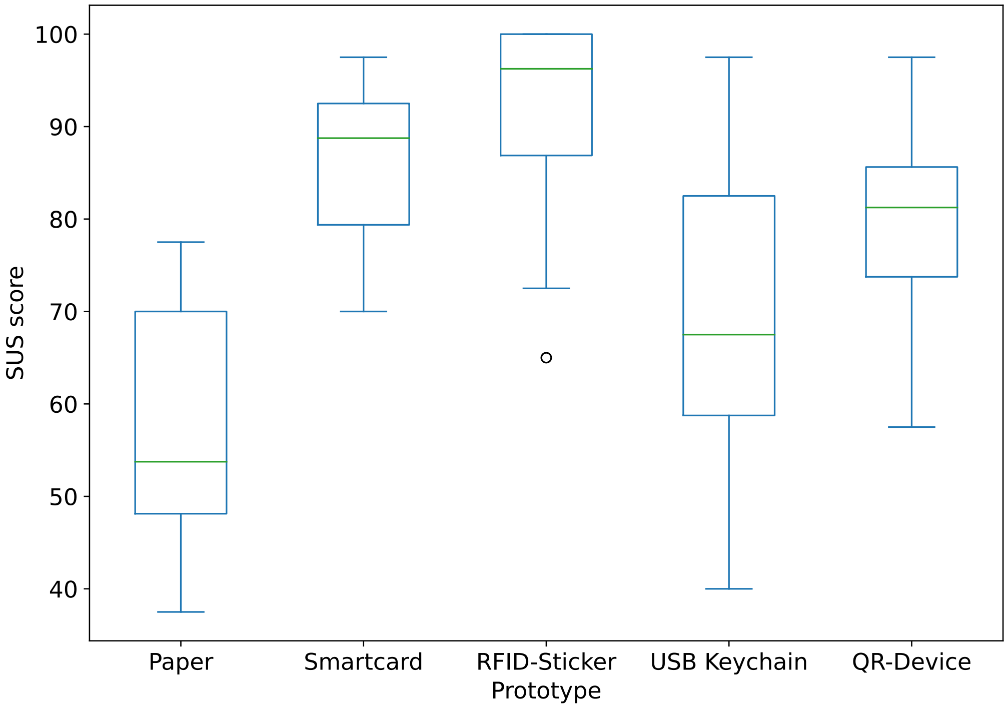


Figure 6.4: SUS scores by prototype, green line is the median.

Prototype SUS Score	Paper	Smartcard	RFID-Sticker	USB Keychain	QR-Device
mean	57.03	86.72	91.72	69.53	78.91
std	13.08	8.50	11.13	17.85	11.03
median	53.75	88.75	96.25	67.50	81.25

Table 6.4: Mean, SD and median of SUS scores per prototype.

Prototype TLX Score	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	38.23	6.61	5.31	15.94	7.29
SD	16.23	5.74	11.38	14.29	6.74
median	33.75	5.83	1.25	14.58	7.50

Table 6.5: Mean, SD and median RTLX scores for each prototype.

Time measurements for key transmission Timing measurements were taken for each task during the second run.

The timing measurements for the QR-Device for task 1 do have a reduced validity. Due to the workstation, where the timestamps have been taken, not being properly synchronized with an ntp server, all timestamps have an offset of around 2980 seconds. This is not an issue for the measurements of smartcard, RFID-Sticker and USB-keychain, as the relative time difference between timestamps for each device is still correct. For the timestamps of the QR-Device, the calculation of time needed for each task becomes more difficult, as the timestamps taken on the QR-Device are properly synced.

To calculate the offset as precisely as possible, we took the timestamps t_{start} of the task start

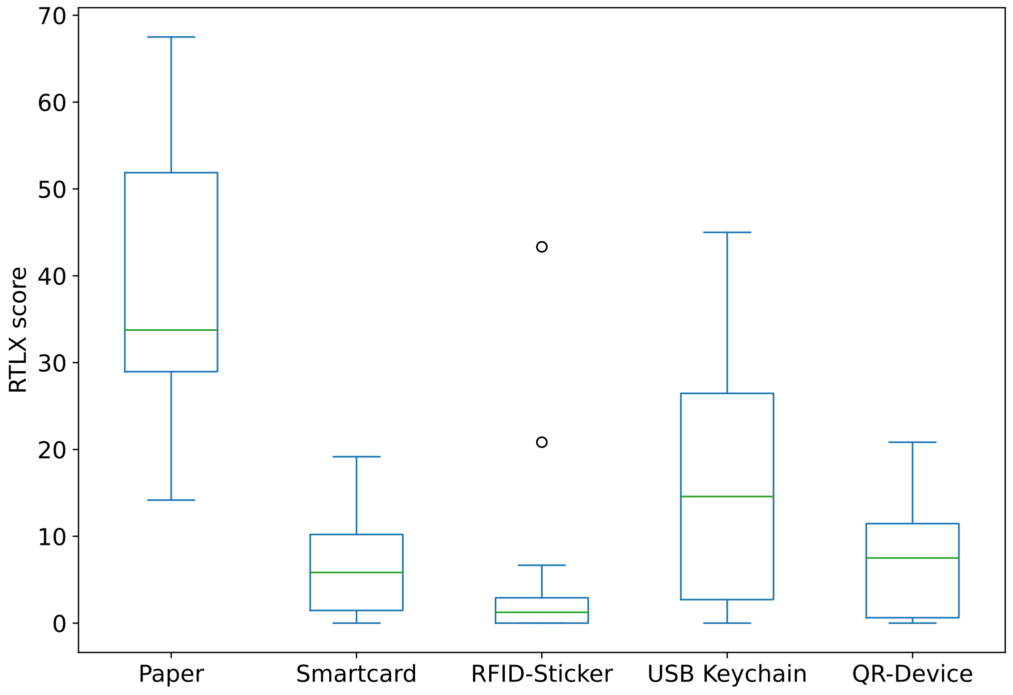


Figure 6.5: Raw (unweighted) TLX scores for each prototype.

event from P_5 and P_8 and determined with the video recording how much time passed until they pressed the interaction complete button on the QR-Device (t_{passed}). The button press timestamp t_{done} was also saved in the QR-Device time logs. Now calculating t_{offset} with $t_{done} - t_{passed} - t_{start}$, resulted in an offset of 2982.056652 seconds for P_5 and an offset of 2981.675639 seconds for P_8. Averaging these two offsets to reduce imprecision of that method, the final offset was 2981.866146 seconds, with which the collected timestamp data for task 1 was corrected (digits after 6th decimal were dropped). This affected only the time measurements for the first nine participants, the other seven time measurements are correct, as the office workstation synchronized its system time after the ninth participant.

The duration for each prototype to read a key is displayed in fig. 6.6. The mean values for each prototype are displayed in tab. 6.6. Additionally the time taken until the participants clicked done and completed the task are displayed in tab. 6.7. Only These values have a low validity, as not each participant clicked the done button, when the interaction was completed. Some participants clicked the done button before e.g. ejecting the USB-keychain, other participants forgot completely to click on ”Done”. At least two times, this was not noticed by the researcher, resulting in a mean time until ”done” was clicked for the QR-Device of 25.39 s (SD: 50.84, median: 6.99). As no action was required for the paper prototype for task 1 a value of ’N/A’ has been inserted in the table.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	N/A	2.74	2.55	14.78	3.93
SD	N/A	0.62	1.42	5.59	1.13
median	N/A	2.69	2.09	14.86	3.79

Table 6.6: Timings for the key transmission during task 1,given in [s], no timings taken for Paper.

While no measurement error occurred during the recording of the timestamps of the second task, it is to be noted that the camera initialization for reading the QR code from the QR-Device,

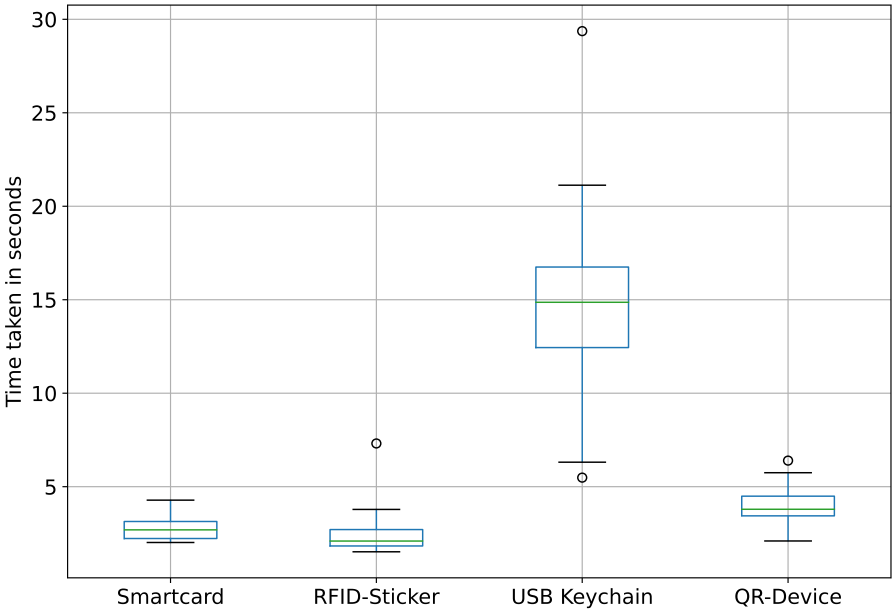


Figure 6.6: Time needed to transfer the key to each prototype.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	N/A	6.81	5.18	22.98	25.39
SD	N/A	1.71	2.30	13.81	50.84
median	N/A	6.80	4.45	19.43	6.99

Table 6.7: Time until participant pressed ”done” button during task 1, all timings given in [s], no timings taken for paper.

takes around five seconds to initialize after clicking the start button. This automatically results in a lower bound of five seconds for time until key transmitted. Some participants used these five seconds to already select a key and hold the device in front of the camera, while others did not. The times needed are depicted in fig. 6.7 and the means, SD and medians of time taken until key transmission and interaction done in tab. 6.8 and tab. 6.9. Directly notable in tab. 6.9 is the mean of 49.11 seconds (SD: 178.19, median 4.35), caused by a participant forgetting to click done and the timer running until the participant finished the questionnaire and starting with the next prototype. As the ”done clicked” timestamp is already imprecise, due to not all participants clicking the done button when intended, the just mentioned outlier was not removed from the dataset.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	89.74	5.19	2.08	10.32	8.59
SD	24.18	6.13	0.94	4.16	5.18
median	80.29	2.93	1.79	10.61	6.73

Table 6.8: Time until keys transmitted during task 2, given in [s].

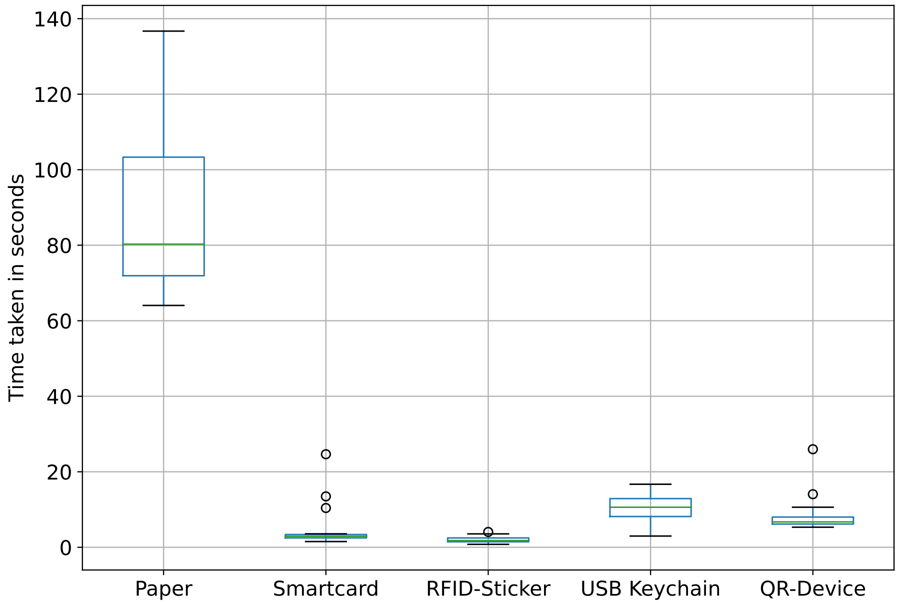


Figure 6.7: Time needed in [s] to transfer key from each prototype to the home office PC.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	94.75	9.27	49.11	15.22	11.09
SD	25.17	6.48	178.19	6.49	5.08
median	85.31	7.21	4.35	14.08	9.23

Table 6.9: Time until participant pressed ”done” button during task 2, times given in [s].

6.3.2 Tangibility

Consequences of loss: The results on our self designed statements on tangibility are presented in this section. For the first statement ”I am aware of the consequences of losing the prototype.” the majority of participants agrees or strongly agrees, regardless of the prototype. The awareness of consequences is highest with the paper prototype, with 13 participants either agreeing or strongly agreeing. This is also visible in the mean score of 4.06 (SD: 0.85, median: 4) in comparison to the other prototypes with mean scores ranging from 3.56 to 3.69.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	4.06	3.69	3.56	3.62	3.62
SD	0.85	1.01	0.96	1.2	1.2
median	4.0	4.0	4.0	4.0	4.0

Table 6.10: Ratings for Tangibility: I am aware of the consequences of losing the prototype.

Feeling in Control: For the statement ”Being able to hold the prototype in my hands, gives me a feeling of being in control”, most participants would agree or strongly agree for the smartcard (N=12), and the lowest number of agreements were observed for the paper prototype with (N=6). The other three prototypes had ten participants (strongly) agree. Half the participants (N=8) would (strongly) disagree with that statement regarding the paper prototype (see fig. 6.9).

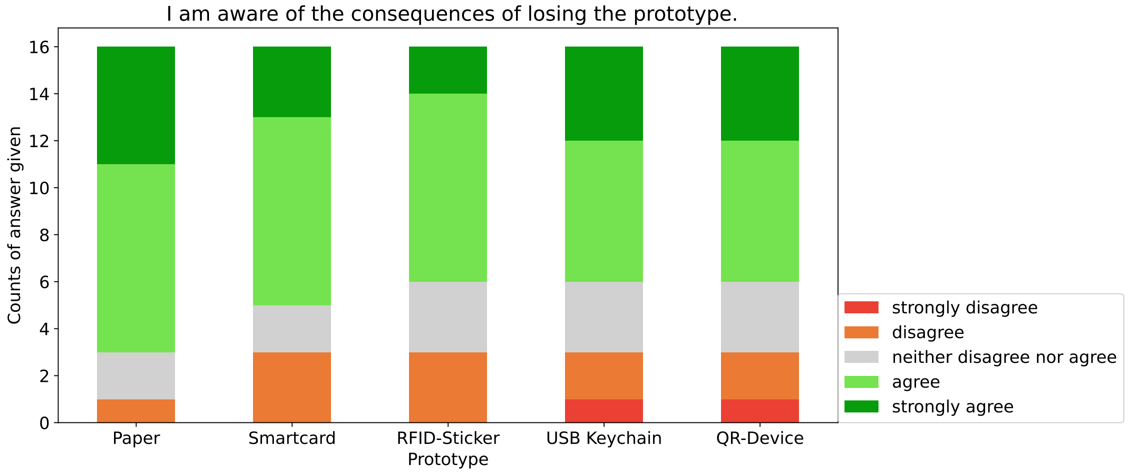


Figure 6.8: Amount of times how often each answer option was chosen for the first tangibility statement.

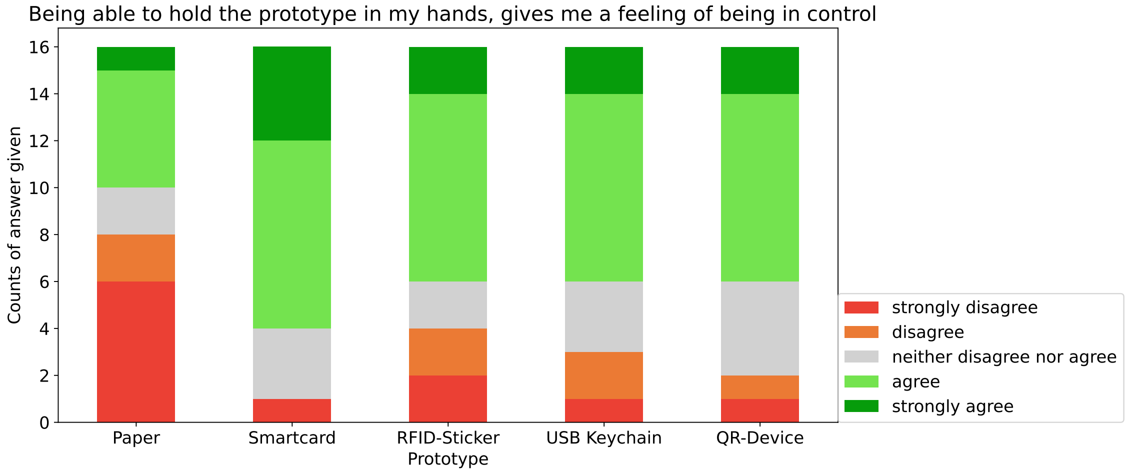


Figure 6.9: Amount of times how often each answer option was chosen for the second tangibility statement.

Preference to a digital alternative When asked about whether a prototype would be preferred to a purely digital alternative, only two prototypes, smartcard and RFID-Sticker, got at least 50% (N=8) of the participants to agree or strongly agree, as depicted in fig. 6.10. With only (N=5) (strongly) agrees the number is significantly lower for the USB-keychain and the QR-Device and with only (N=4) statements of agreement is lowest for paper. The means, SD and medians for this statement are shown in tab. 6.12.

6.3.3 Form Factor

Easily lost?: The highest level of agreement for the statement ”I believe this prototype could get easily lost” can be observed with paper and the USB-keychain, both with N=10 participants agreeing or strongly agreeing. The RFID-Sticker and QR-Device have a smaller risk of getting lost, both with N=6 participants agreeing. The smartcard fares best here, with only N=5 votes on being easily lost, see fig. 6.11. The same can be seen in tab. 6.13.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.56	3.88	3.38	3.5	3.56
SD	1.46	1.02	1.26	1.1	1.03
median	2.5	4.0	4.0	4.0	4.0

Table 6.11: Ratings for Tangibility: Being able to hold the prototype in my hands, gives me a feeling of being in control.

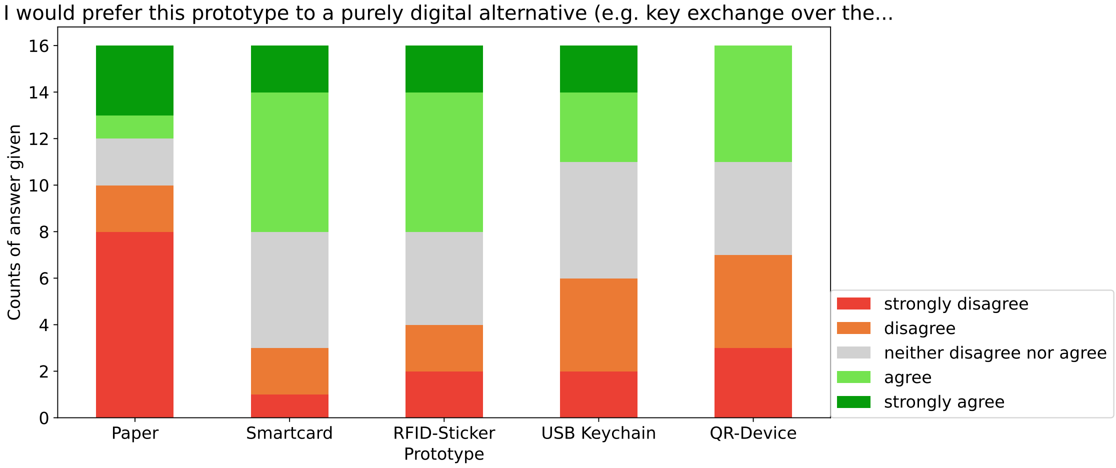


Figure 6.10: Amount of times how often each answer option was chosen for the third tangibility statement.

Easily transportable? For the statement " I believe this prototype is easily transportable" all five prototypes had a majority of the participants agree. The agreement was highest with the smartcard, since all (N=16) participants (strongly) agreed. A close second are the RFID-Sticker and the USB-keychain, both with (N=15) participants agreeing or strongly agreeing. The USB-keychain achieves a slightly better rating with N=13 strongly agrees and no disagree, compared to N=10 strongly agree and 1 strongly disagree for the RFID-Sticker. Paper and QR-Device still get good results, with N=12 (strongly) agreements for paper and N=10 for the QR Device. The QR-Device has the highest amount of disagreements, with N=5 (strongly) disagree. This is depicted in fig. 6.12, means, SD and median shown in tab. 6.14.

Handyness With the exception of paper, the majority of participants rated all prototypes as handy, in agreement with the statement "I believe this prototype is handy". Notable here are smartcard and RFID-Sticker, with both not even receiving one disagree vote and the RFID-Sticker only receiving ratings of (strongly) agree (see fig. 6.13). USB-keychain and the QR-Device receiving ten, respectively eleven (strongly) agrees, while paper comes in last with just 4 ratings of

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.31	3.38	3.25	2.94	2.69
SD	1.62	1.09	1.24	1.24	1.14
median	1.5	3.5	3.5	3.0	3.0

Table 6.12: Ratings for Tangibility: I would prefer this prototype to a purely digital alternative (e.g. key exchange over the internet).

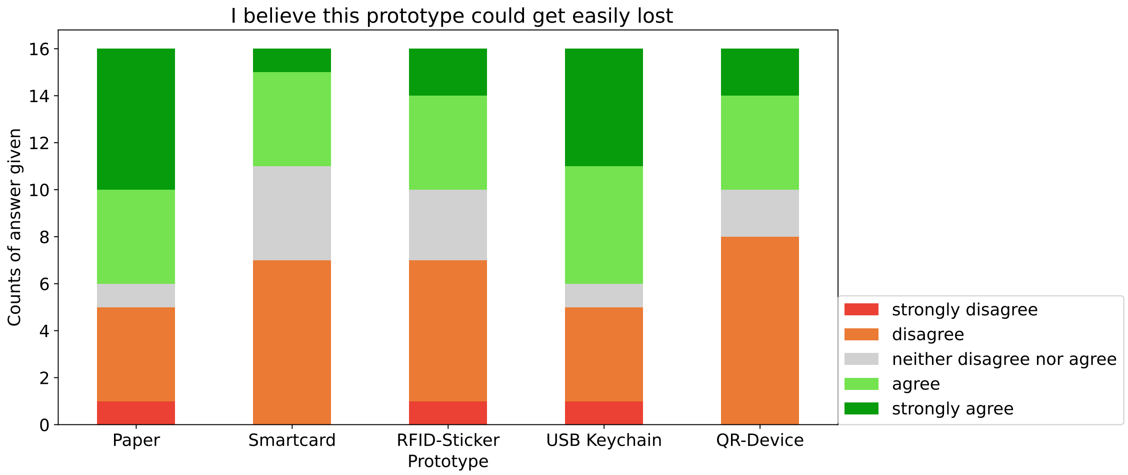


Figure 6.11: Amount of times how often each answer option was chosen for the first formfactor statement. Even though the USB was attached to the users key chain, it is quite easily lost. Prototypes that can placed in (smartcard) or on (sticker) wallets or phonecases are not.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	3.62	2.94	3.0	3.56	3.0
SD	1.41	1.0	1.21	1.36	1.15
median	4.0	3.0	3.0	4.0	2.5

Table 6.13: Ratings for form factor: I believe this prototype could get easily lost.

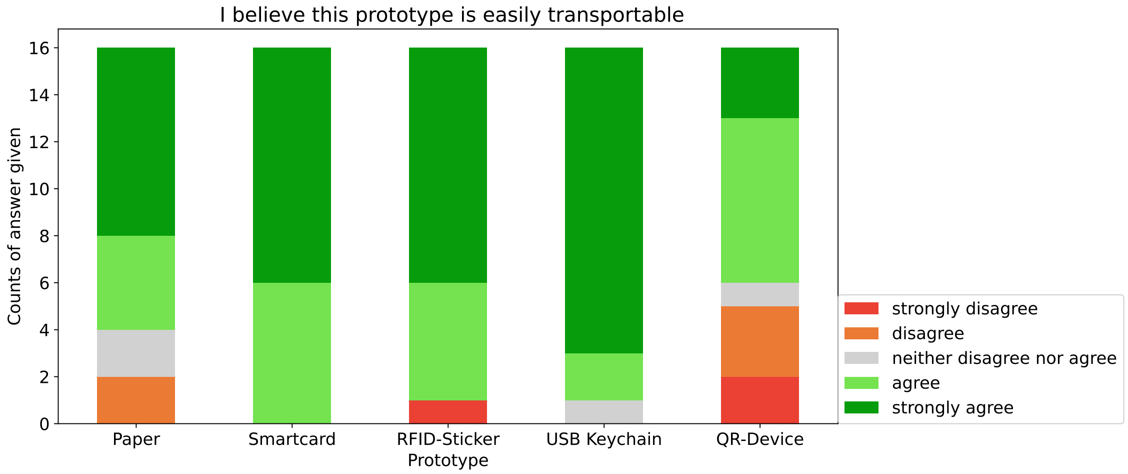


Figure 6.12: Amount of times how often each answer option was chosen for the second formfactor statement. Smaller (or lighter and foldable) prototypes are clearly easier to transport.

agreement. This is also visible in the mean ratings of each prototype, displayed in tab. 6.15.

Fun The last statement regarding the form factor of the device was ”I believe using this prototype is fun”. While the smartcard and the RFID-Sticker, as well as the QR-Device, received a majority of (strongly) agree, paper received only negative feedback with (N=16) ratings of disagree or strongly disagree. USB-keychain placing in the middle with receiving (N=7) disagreements and (N=8) neutral ratings. The means, SD and medians can be found in tab. 6.16.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	4.12	4.62	4.44	4.75	3.38
SD	1.09	0.5	1.03	0.58	1.36
median	1.5	5.0	5.0	5.0	4.0

Table 6.14: Ratings for form factor: I believe this prototype is easily transportable.

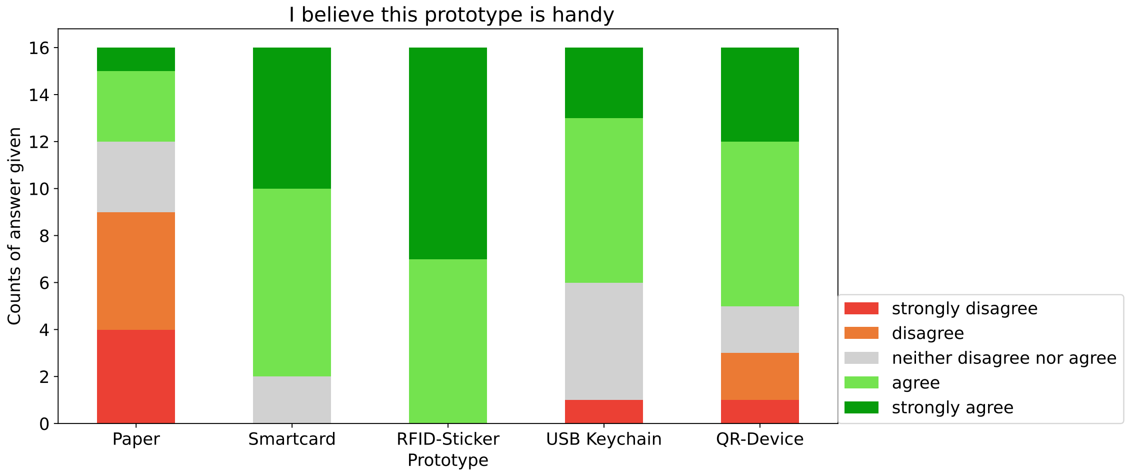


Figure 6.13: Amount of times how often each answer option was chosen for the third form factor statement. The easier the required user interaction, the better ratings the prototype got.

6.3.4 Security

Consequence of unauthorized access The statement ”I am aware of the consequences if another person gains access to the prototype” is among the few statements, in which paper does not fare much worse than the other prototypes. Even more than that, the ratings for each prototype for this statement are nearly identical, showing less difference than the results for the ”consequence of loss” tangibility statement. Visible in fig. 6.15 are the answer distributions, with a difference of at most 1 vote for (strongly) agree or disagree between the prototypes. This is also visible in tab. 6.17, with mean values for all prototypes between 3.75 and 4 and a median of 4 for all prototypes.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.5	4.25	4.56	3.69	3.69
SD	1.26	0.68	0.51	1.01	1.2
median	2.0	4.0	5.0	4.0	4.0

Table 6.15: Ratings for form factor: I believe this prototype is handy.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	1.06	3.81	4.06	2.44	3.38
SD	0.25	0.98	1.18	0.89	1.41
median	1.0	4.0	4.5	3.0	4.0

Table 6.16: Ratings for form factor: I believe using this prototype is fun.

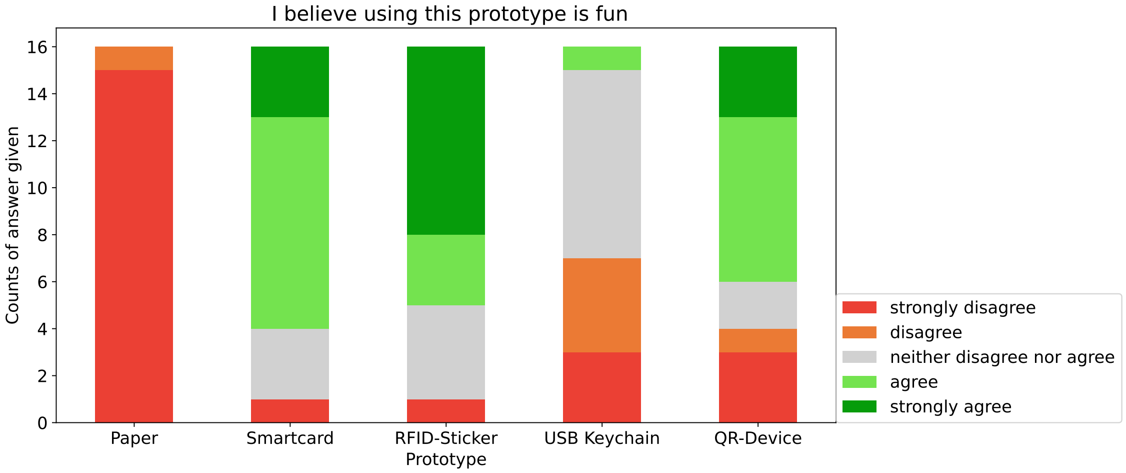


Figure 6.14: Amount of times each answer option was chosen for the fourth form factor statement.

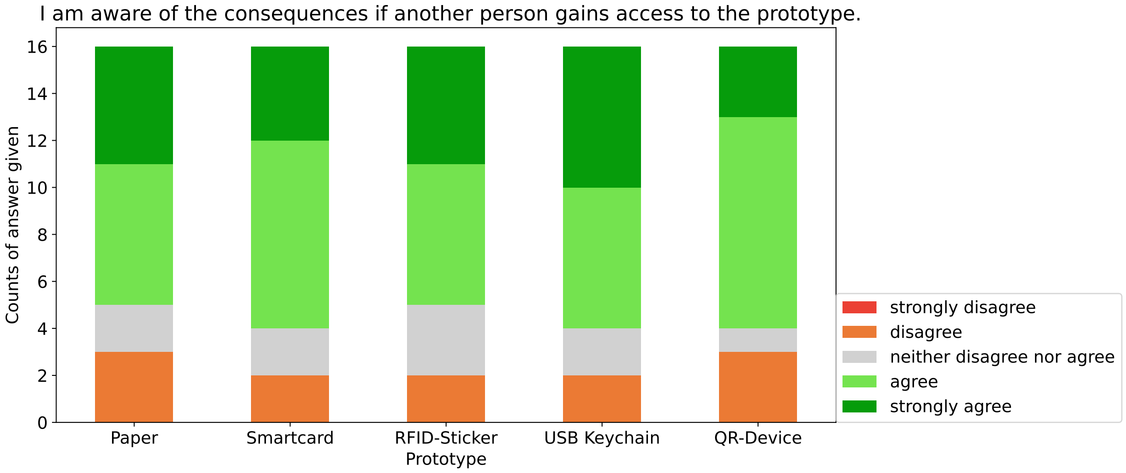


Figure 6.15: Amount of times how often each answer option was chosen for the first security statement.

Secure storage For the statement ”I believe the prototype is keeping the keys secure” only the smartcard and the QR-Device received a higher number of (strongly) agree rating with (N=11) for the smartcard and (N=9) for the QR-Device. The other 3 prototypes received a similar amount of (strongly) agrees, between (N=4) and (N=6). Both the USB-keychain as well as the paper received a larger amount of (strongly) disagrees, (N=8) for paper and (N=6) for the USB. The RFID notably received a larger amount (N=9) of neither agree nor disagrees, depicted in fig. 6.16, the mean, SD and medians can be found in tab. 6.18.

Secure transmission For the statement ”I feel the transmission of the keys to and from the prototype is secure”, the smartcard received the best results with (N=10) (strongly) agrees and no disagrees. The reason here could be the same, as for the statement of securely storing the keys, that participants implicitly trust the smartcard and the (simulated) physical contact required for the transmission of the keys, as this is a interaction with a device the participants are already accustomed to. Both the USB and the RFID sticker have similar results for that statement, with the same amount (N=6) of (strongly) agrees an (N=2) (strongly) disagrees. The disagree statements for the USB are strongly disagree, while also having (N=2) votes of strongly agree. The RFID sticker in comparison has no votes on the extreme ends of the scale, as can be seen in fig. 6.17. A

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	3.81	3.88	3.88	4.0	3.75
SD	1.11	0.96	1.02	1.03	1.0
median	4.0	4.0	4.0	4.0	4.0

Table 6.17: Ratings for Security: I am aware of the consequences if another person gains access to the prototype.

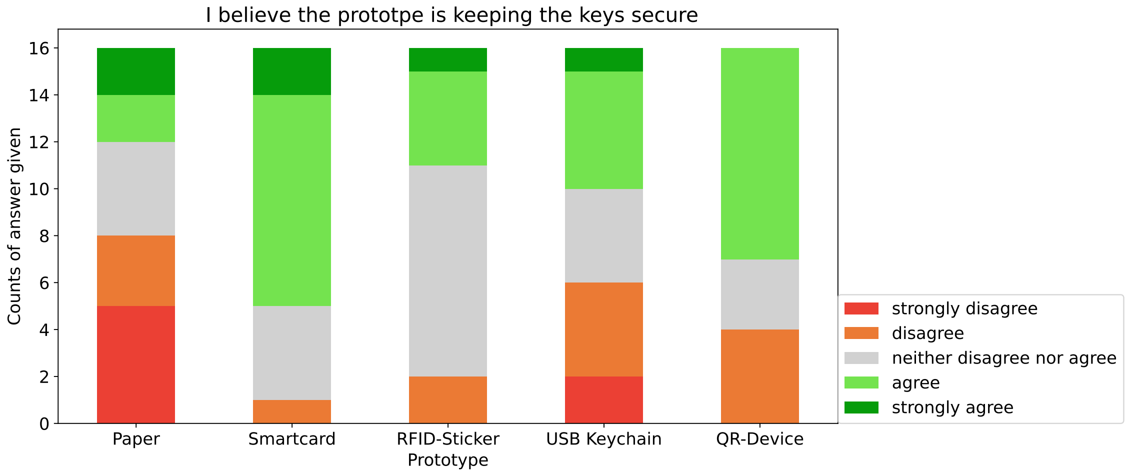


Figure 6.16: Amount of times how often each answer option was chosen for the second security statement.

close second to the smartcard is the QR-Device, which can be seen when looking at the mean and median scores in tab. 6.19, as both devices have the same median rating of 4, with the smartcard having a higher mean of 3.88 (SD:0.81) compared to the QR-Devices 3.5 (SD: 0.97).

6.3.5 HCTS

For the evalution of the HCTS questionnaire, the same method was used as in [35]. For the overall HCTS score the values of the answer options were summed and as the items for perceived trust were negatively worded, so for the determination of the final score the values were inverted.

Perceived Risk Perceived risk is highest for the USB and paper, as can be seen in fig 6.18, mean, SD and median can be seen in tab 6.20.

Benevolence All devices archieve similarily low values (ratings less than 3, visisble in tab 6.21) on the benevolence subscale, with the exception of paper, which receives the worst ratings of all in this category. The results are visualized in fig 6.19 and may be explained by the way the statements for this subscale are worded.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.56	3.75	3.25	2.94	3.31
SD	1.41	0.77	0.77	1.18	0.87
median	2.5	4.0	3.0	3.0	4.0

Table 6.18: Ratings for Security: I believe the prototype is keeping the keys secure.

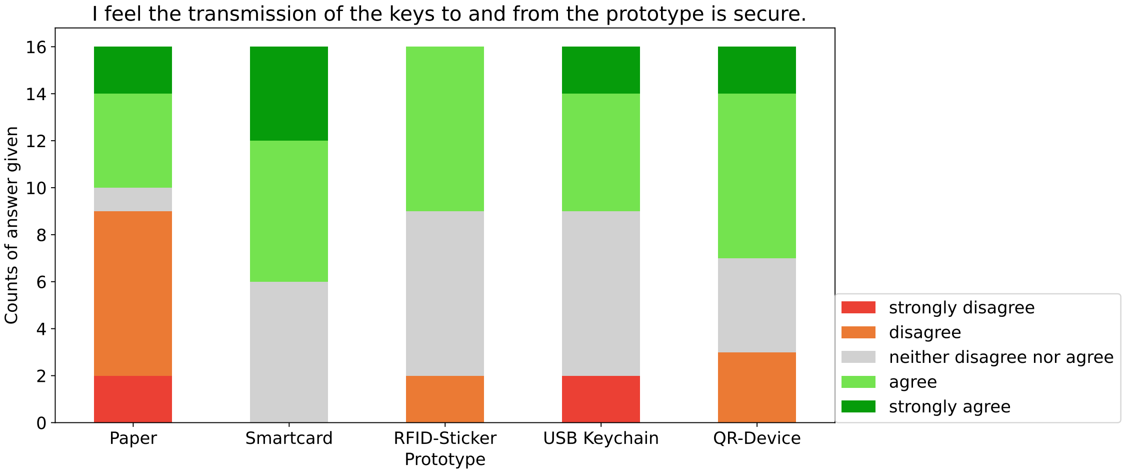


Figure 6.17: Amount of times how often each answer option was chosen for the third security statement.

Prototype	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.81	3.88	3.31	3.31	3.5
SD	1.33	0.81	0.7	1.14	0.97
median	2.0	4.0	3.0	3.0	4.0

Table 6.19: Ratings for Security: I feel the transmission of the keys to and from the prototype is secure.

Competence For competence, the values between each prototype (with the exception of paper) do not differ that much, as the means for smartcard, sticker and USB are all between 3.54 and 3.6, as depicted in tab. 6.22. Only the QR device has a slightly higher mean value, which gets put into perspective when considering the median values of the better four prototypes. This becomes apparent, when looking at fig. 6.20, in which the boxes between the prototypes (not counting paper) look quite similar.

Trust The mean values of trust for each prototype rank the smartcard as the highest, followed by the QR-Device, as shown in tab. 6.23. The ratings for trust (depicted in fig. 6.21) are being close to the inverse of the ratings of perceived risk.

Overall HCTS score The overall HCTS scores for each prototype are shown in ta. 6.24 and visualized in fig. 6.22. As the overall score is calculated from the same items of the subscales, it is not surprising to see the same high overall trust score for the smartcard and the QR-device closely followed by the RFID-Sticker and the USB. Even though no prototype achieved very high results

	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	3.60	2.33	2.69	3.04	2.73
SD	0.79	0.81	0.66	0.76	0.90
median	3.67	2.33	2.67	3.17	3.00

Table 6.20: Mean,SD and median scores for the perceived risk subscale of HCTS (item 1-3), rated on a scale of 1-5, lower values are better.

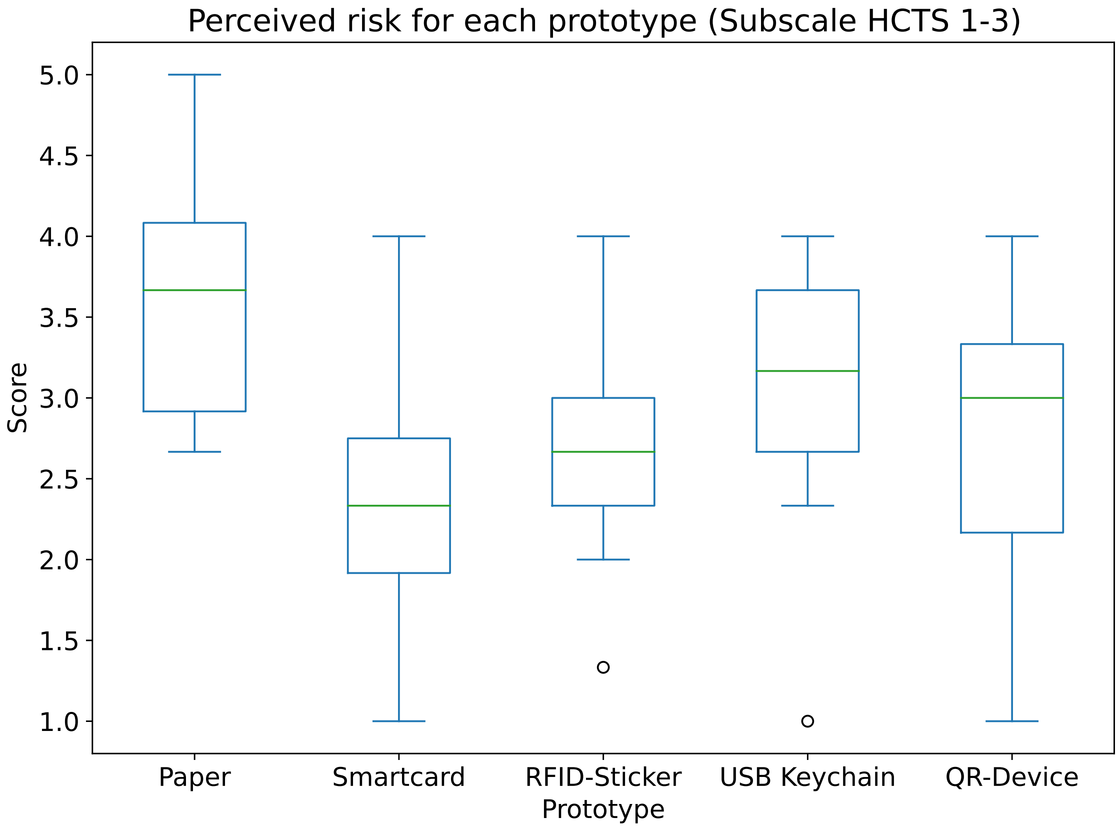


Figure 6.18: Results of perceived risk for each prototype, lower values are better.

	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	1.98	2.85	2.69	2.52	2.73
SD	0.79	0.64	0.72	0.80	0.65
median	1.83	3.00	2.83	2.67	3.00

Table 6.21: Mean,SD and median scores for the benevolence subscale of HCTS (4-6), rated on a scale of 1-5, higher values are better.

all (with the exception of paper) reach at least a medium level of trust, with the USB and sticker having the lowest values.

6.3.6 Usage

The results on what category the participants would use a prototype for will be presented in this section. The usage categories can be sorted into categories based on the level of sensitivity of the data accessible in each category. Under ”highly sensitive” would be critical infrastructure and online banking which is one of the most sensitive types of information for Germans [38]. In this context, work related data, work email and personal email is considered ”medium sensitive” and everything else (streaming service, social media and the general entertainment category) are considered ”low sensitivity” data. Even though social media profiles are seen with a medium level of sensitivity compared to other personal information [38] , compared to work data and online banking data, we can categorize it as low here. The usage numbers are shown in tab. 6.25 as well as graphically represented in fig. 6.23.

Due to the way the question was worded, it was not specified, if the usage was the same as

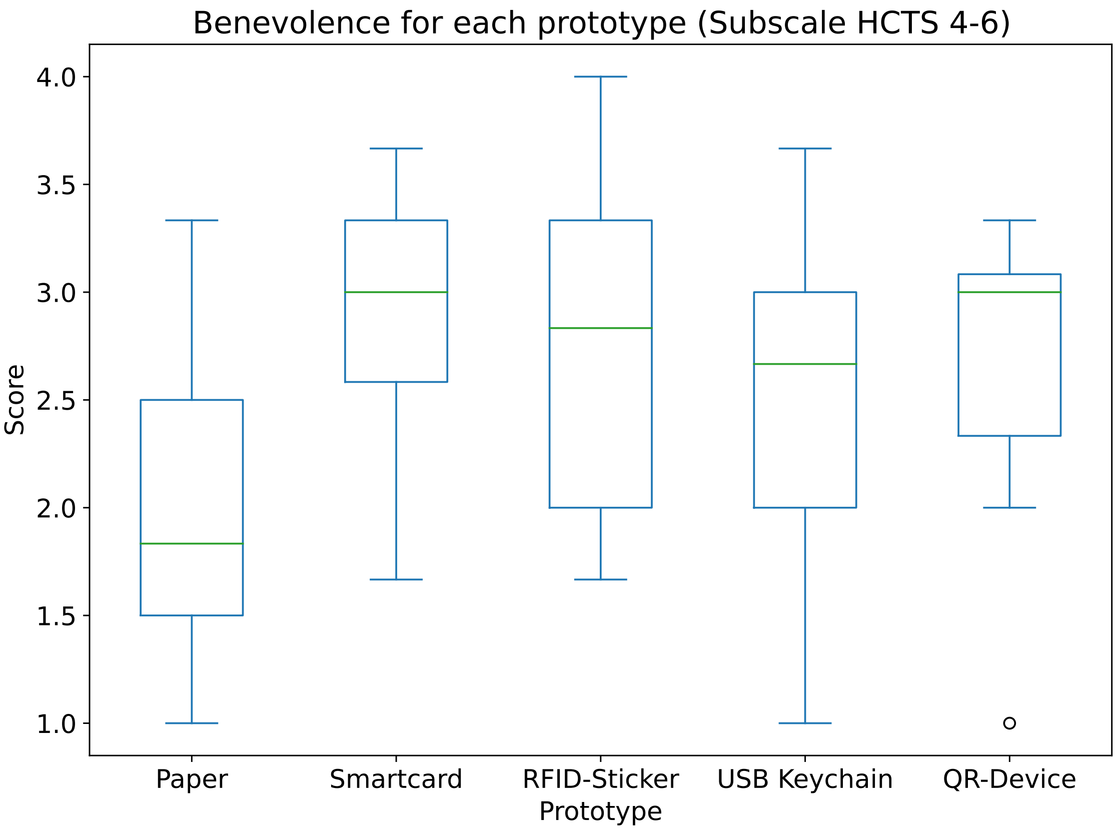


Figure 6.19: Results of the aspect of benevolence for each prototype.

	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.23	3.60	3.56	3.54	3.79
SD	1.21	0.96	0.80	0.97	0.79
median	2.00	4.00	3.67	3.50	3.83

Table 6.22: Mean,SD and median scores for the competence subscale of HCTS (7-9), rated on a scale of 1-5.

in the use case, e.g. one time or at least rarely the transport of a key, or if it would take the role of an authentication device, which would require using it every day to login to work systems. Unfortunately, this is something that can only be speculated on.

6.4 Qualitative Analysis

The main focus of this section is the analysis of the participants interview and other information that can be gained from the video recordings of the study. Not all participants mentioned all devices, as the questions were worded in a way, that allowed the participants to decide which devices they want to comment on in regard to each question.

6.4.1 Usability

The first and third interview question directly targeted aspects of usability in day to day life and improvements in usability. As seen during the quantitative analysis, the paper prototype got bad ratings for usability, a trend that can be observed in the participants interviews, too.

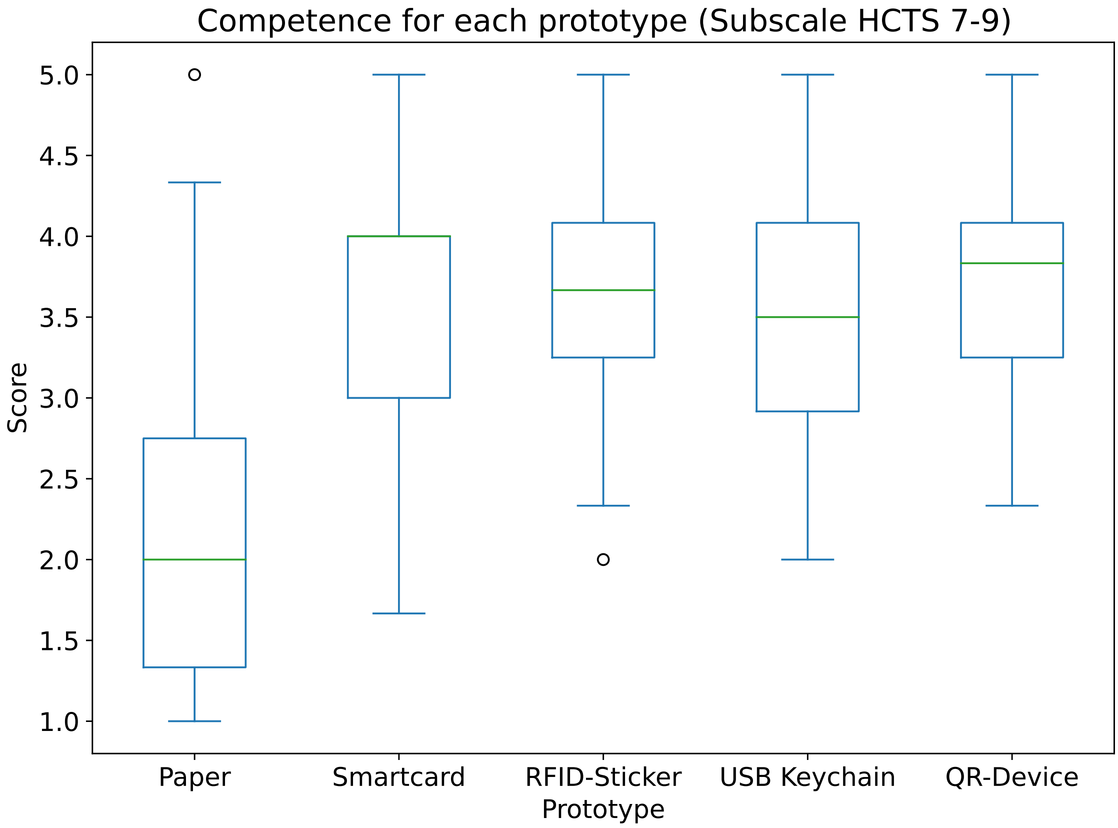


Figure 6.20: Results for the aspect of competence for each prototype.

	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	2.73	3.58	3.25	3.25	3.40
SD	1.02	0.80	1.01	0.90	0.67
median	2.67	3.67	3.50	3.33	3.33

Table 6.23: Mean,SD and median scores for the trust subscale of HCTS (10-12), rated on a scale of 1-5.

Strong statements as ”[...] paper, throw it away and use something else” (P_7) capture the results of the paper prototype. Some participants did not want to use the paper prototype at all (P_4, P_7) and P_1 only ”if i have to”. Paper was generally mentioned as cumbersome (N = 5) with the key entry method being difficult to use (N = 8). Some participants explicitly found the key too long or would want a shorter key (N = 3) , one would accept more complexity in the form of special characters (P_{13}).

Recommendations for improving the usability of the paper prototype focused on the formatting of the printed key (N = 3) or the UI of the app, where the key had to be entered (N = 2). P_{10} suggested to have the user memorize the key, so an additional key transport device would not be needed, but that same participant did not want to use any key transport device at all.

In stark comparison to paper are the smartcard and the RFID-Sticker, with both receiving nearly no criticism. Only one participant (P_{14}) found the smartcard use slightly cumbersome, due to the handling of the card reader. The other participants rated the smartcard as easy usable (N = 5) and fun (N = 2). The ease of use may relate to participants being already accustomed to the use of smartcards e.g. for banking, mentioned by (N = 2) participants. Some problems occurred though, as a few (N = 2) participants did not insert the card completely or inserted it too fast during

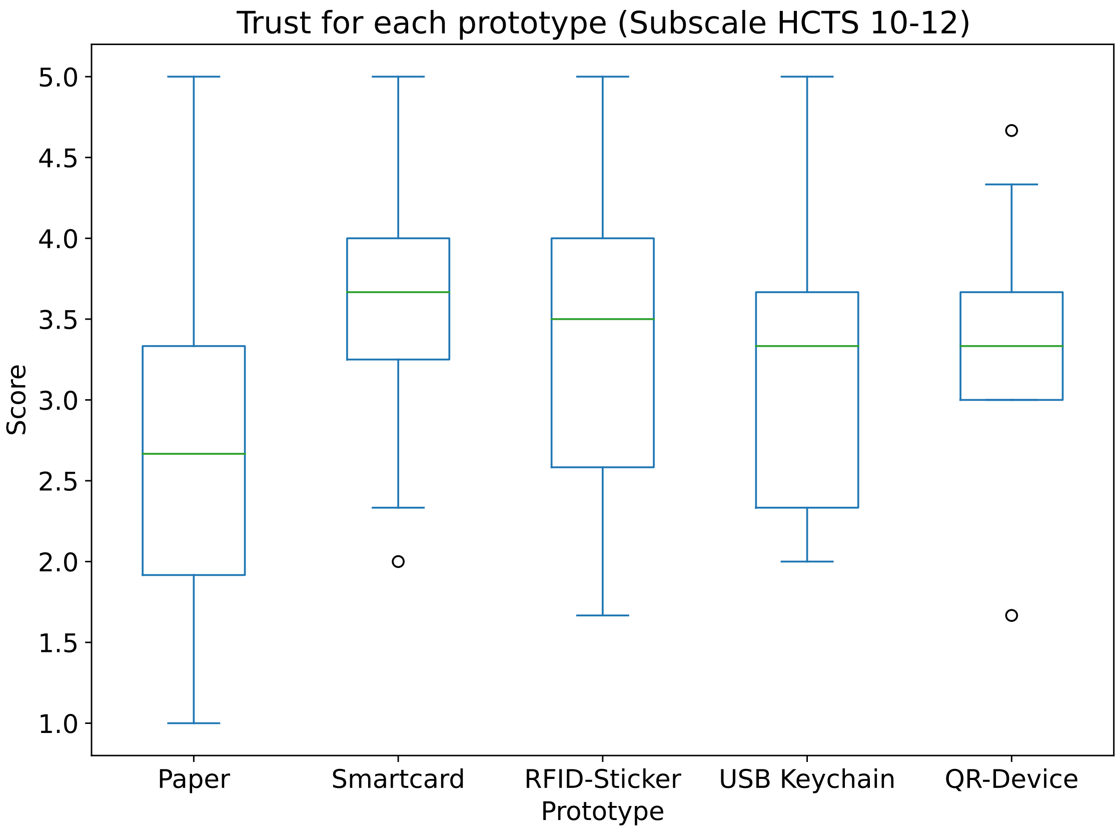


Figure 6.21: Results for the aspect of trust for each prototype.

	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
mean	28.00	41.12	38.44	36.81	39.56
SD	7.27	6.38	6.40	7.46	6.48
median	26.50	42.00	40.50	37.00	39.50

Table 6.24: Mean,SD and median scores for overall HCTS score, higher values are better.

the second task. This led to errors of the card not being read. As this occurred during the test run for the prototype, the participants then knew for the second run that these errors could occur and paid attention to insert the card correctly. These errors are not specific errors in the use of the prototype, but could also be caused by the RFID module itself, as their build quality can vary quite a bit. Even better than the smartcard were the answers for the RFID-Sticker with (N = 6) rating it as easy usable and P_6 additionally rating it fun.

The USB prototype only got few mentions for usability during the interview. P_1 gave the same answers as for paper and would only use it if they had to. P_8 thought the device was easy to use and liked the option of saving multiple keys. An other participant (P_4) found the prototype by itself easy to use, even if was slightly cumbersome and slow. P_7 would only use the USB because of its easy replaceability. To improve the usability, a suggestion that was made by (N = 2) participants was to have the key save or load happen automatically, without the user having to use the file dialog.

As the QR-Device placed in the middle between the best and average devices, the amount of usability comments is low. It was found easy (N = 2) or at least ok (N = 1) to use, with the only usability feedback focusing on improving the UI of the QR-Device (N = 3).

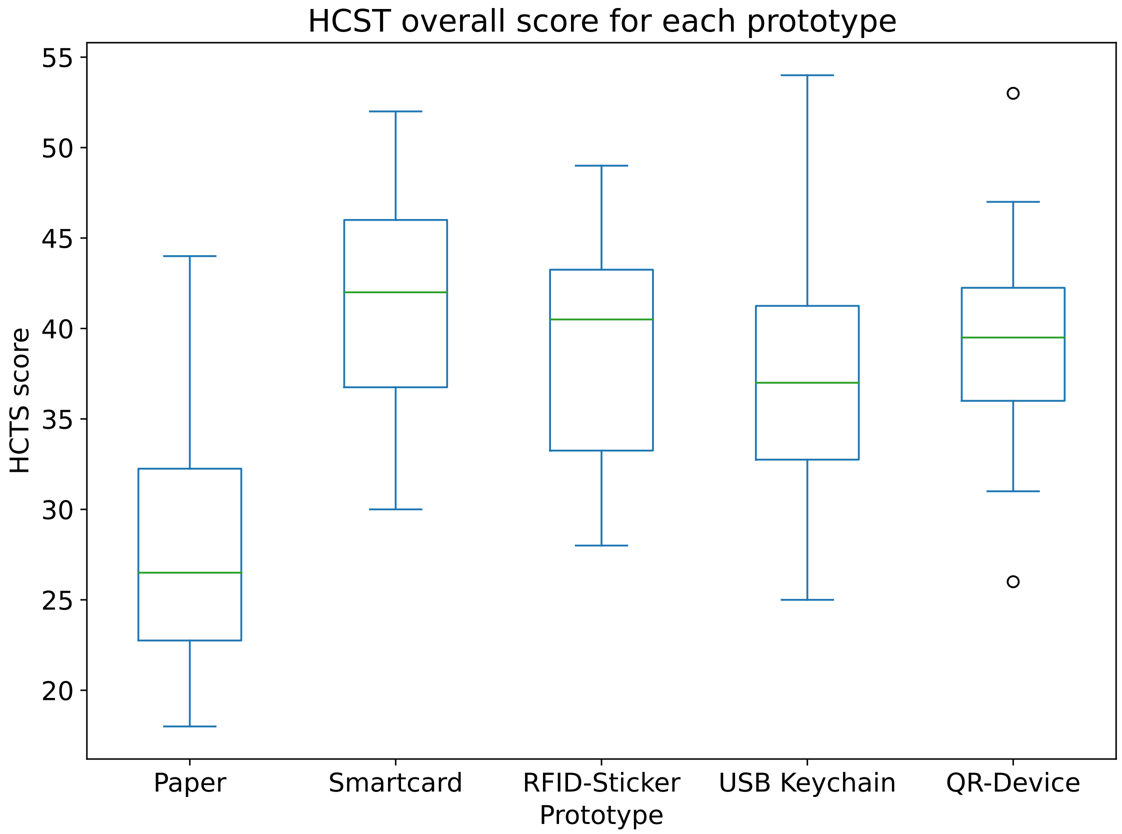


Figure 6.22: Overall HCTS scores for each prototype, from 12 (very low) to 60 (very high).

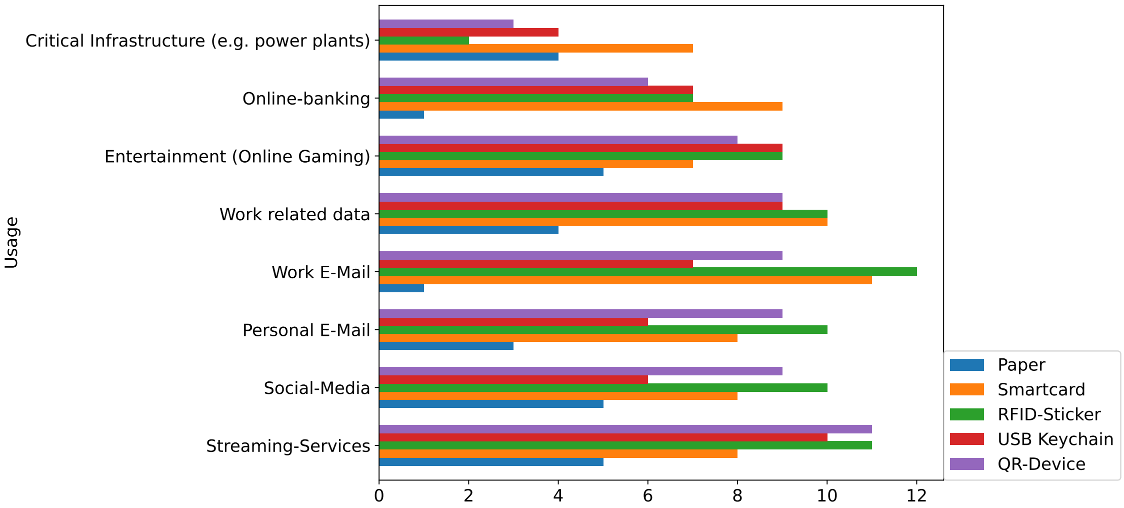


Figure 6.23: Number of participants willing to use each prototype for each category.

6.4.2 Tangibility

During the interview, various aspects of each prototype were mentioned in relation to physical aspects of each prototype. These are based primarily on the question if it would bother the participants to transport the prototype.

The opinions on the paper prototype were divided. Some participants (P_1, P_4) were bothered by transporting it, while P_2 and P_{10} thought it was easy to transport, as it can be folded and carried inside a wallet or pant pockets. Multiple participants (N = 3) feared that the paper could easily be

Prototype Usage category	Paper	Smartcard	RFID-Sticker	USB keychain	QR-Device
Streaming-Services	5	8	11	10	11
Social-Media	5	8	10	6	9
Personal E-Mail	3	8	10	6	9
Work E-Mail	1	11	12	7	9
Work related data	4	10	10	9	9
Entertainment (On...	5	7	9	9	8
Online-banking	1	9	7	7	6
Critical Infrastr...	4	7	2	4	3

Table 6.25: Number of participants willing to use a prototype for a specific usage category. The absence of a 'don't want to use' option, forced each participant to select at least one category they could imagine using the prototype for.

lost, while P_{10} found it hardest to loose compared to the other prototypes. The most common fear for paper was it being easily damaged, be it by tearing or water damage (N = 3).

The smartcard instead only got the feedback, that it is quite easy to transport, as most people already carry some cards around (N = 5).

The sticker was as easy to use as the smartcard, but got more feedback on its physical characteristics. P_{15} said the sticker was easy to transport, P_{10} depending on the object it is placed on. Multiple participants (N=5) were "concerned, that [the sticker] comes off or gets lost"(P_{11}). In that same line is that (N=3) participant would prefer the sticker being integrated in a larger physical object e.g. a smartcard, which was mentioned as a potential improvement. An interesting aspect here is to evaluate the objects chosen by the participants to place the RFID-Sticker on. Most participants (N=11) decided to place the sticker on their smartphone. As one mentioned, this allows it to be easily located as the phone can be called (P_8) and most people carry their phones with them all the time. For (N=2) participants the object of choice was their wallet, others chose their e-cigarette case(P_{14}), their watch (P_{11}) or a random piece of paper (P_{16}). P_{16} mentioned after the study, they chose paper, as they did not know how the RFID-Sticker worked and feared that interference on e.g. on the case of their headphones would not allow the sticker to work properly. The most interesting case here is P_{11}, as placing the sticker on the watch basically transforms the sticker from a key carry object into a wearable, which in turn changes the user interaction, as can be seen in fig. 6.24.

For the USB-Keychain P_5 mentioned it being secure, because due to the attachment "it can't come lose so easily". This is in contrast to (N=3) participants that feared the USB drive could easily get lost, with P_{10} specifying that the string used to attach the USB breaking quite easily and instead use a USB that can hook directly into a key ring. An other potential improvement mentioned (N = 3) for the USB is using one with a bigger housing or a bigger USB drive in general. P_{14} suggested the use of a USB-C type connector to make the use easier.

Many participants (N=6) mentioned the QR-Device being too big or being bothered to transport it (N=2) because they had to carry an additional device (N=4). P_6 explicitly said the QR-Device is too heavy. One improvement could be to decrease the size to just be able to fit a set of keys (P_4).

A general opinion voiced was that all the smaller devices are easier transported (N=3).

6.4.3 Security

For the security of the paper prototype, the most (N=7) mentioned aspects was it being insecure due to the key being printed in plaintext and therefore easily accessible by anyone. Not having

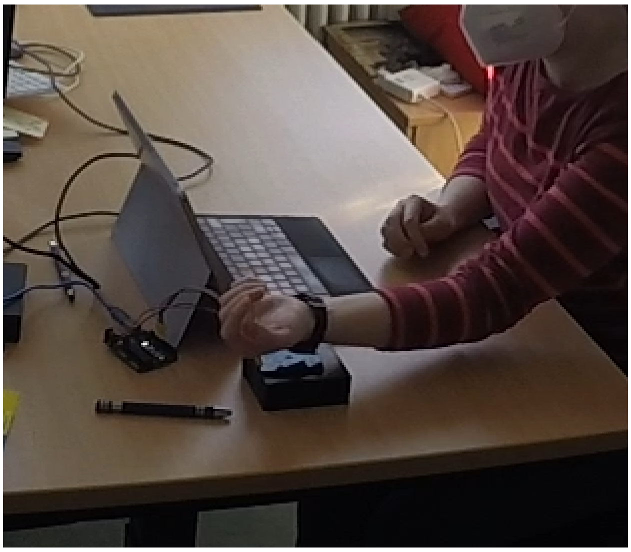


Figure 6.24: A participant transforming the sticker into a wearable key transport device by attaching it to their watch, therefore changing the user interaction for the transmission of keys.

any digital components was seen as making the device secure (N=2), while (P_6) felt the opposite as ”analog somehow [...] looks the least secure”.

Contrasting paper, the smartcard was perceived as secure (N=5) even though one participant mentioned that it may be possible to just read the card contents when in close proximity (P_{15}). To improve the security of the smartcard, the use of an additional PIN was mentioned (N=2). P_{15} mentioned the use of RFID blocking transport option, as some wallets already offer, to increase the security. This is only necessary, if the card actually has any contactless functionality.

The proximity unauthorized reading issue was also mentioned as reason for the RFID-Sticker being insecure (N=6). Without mentioning their reason P_{11} perceived the sticker as insecure, while P_6 and P_8 thought the sticker is secure. An improvement for the RFID-Sticker mentioned by P_8, was something that was already mentioned by a participant in the focus group, RFID chip as a subdermal implant. As with the smartcard, the use of a signal blocking case to improve the devices security was mentioned by P_2.

Most times the USB-Keychain was mentioned in a security context, it was done positively. P_2 thought the USB secure as ”someone would need to get it in their hands and plug it in somewhere”, thus need direct physical access to get the keys. If the USB is only used for the transport of the keys and nothing else P_9 thought the device to be secure. P_{11} felt the USB to be secure as they could be sure the keys were only saved locally.

Compared to the USB more participants thought the QR-Device to be secure (N = 3), with P_7 specifying that this is because it is a dedicated device. P_2 thought the QR-Device being insecure, as the time when the QR code is displayed on the key issuing device could be vulnerable to a shoulder surfing attack photographing the QR code. P_{11} mentioned privacy concerns as the reason of the device being insecure, because of the way the laptop cameras preview of the participant during task 2 and what could happen if an attacker were to gain control over the camera. Of course, if an attacker actually managed to gain control over the camera, it is highly likely that the whole system is compromised and no security guarantees on that system would hold up. An interesting concept was proposed by P_2 to increase the security (and reduce the risk of shoulder surfing attacks), is dynamic positioning of the QR code on the screen. In that concept, the position where the QR code will shop up on the screen is shown to the user before, so they can already position the QR device in read mode correctly, and then the QR code is only displayed for the smallest necessary of time for the QR-device to read the code. A more common improvement of security would be the adding of a PIN or similar to the QR code to prevent unauthorized access, proposed by (N=2) participants.

Many more general security improvements were proposed, that could be applicable to all (or at least more than one device). P_6 suggested a general addition of a 2nd factor to authenticate the user as being authorized to transfer the keys, which is in line with a adding PINs or passwords to improve overall security mentioned by (N=3) participants. P_8 had two interesting improvements, as they already mentioned the sticker on the smartphone being easily locateable, to add the option to other devices as well. While this could be challenging from a technical perspective it may reduce the risk of losing the devices. The other idea brought forth was the option of keys autodeleting if not used in some specific timeframe, but this could better be implemented in the overall design of the system, that issued keys have a time limited validity until the have to be first used, similar to some password reset links for various internet services.

The last aspect mentioned to improve security was user education, where (N=2) participants thought this important enough to mention. This would not only contain explanations on the technical functionality of the devices,as some participants did not want to provide a assessment of the perceived security, due to not knowing how the devices worked (N=2).

When asked the last question, if additional security features like a PIN or fingerprint would make a difference, (N=7) participants said that additional features would make the devices feel more secure, with (N=4) participants directly referring to PINs and (N=3) participants referring to fingerprint as the feature that would make the device more secure.

6.5 Summary

This section will summarize the results and answer the research questions established in 1.2.

Overall only two devices can be seen as having performed reasonably well overall, the smartcard and the RFID sticker. The results for the paper prototype were as expected, due to forcing the users to type in a 64 character random sequence. This is already munitioned in previous works, having to type in a random sequence of characters is generally a bad user experience, as it is slow and error prone [36]. With this in mind, the rest of the summary and discussion will focus on the remaining four prototypes, that each offer a unique form of user interaction for the transmission and storage of cryptographic keys.

The first focus will be on the usability of the prototypes. To evaluate the overall usability, we take the direct SUS scores into account, the RTLX scores of how difficult it was to perform the required tasks with the prototype, as well as the time required to complete each task

The second best rated device (smartcard) has received great ratings for usability, when SUS score, TLX results and the time needed for the storing and retrieving of the keys is taken into account. In the aspects for form factor and tangibility, the smartcard also performed well and barely any negative feedback was given about the smartcard. The smartcard also has a high level of perceived security with the best results for key storage and transmission, resulting in the highest overall trust score and the lowest level perceived risk, making it overall the best device of the ones evaluated.

A close contender is the RFID-Sticker, which while not faring as good in perceived security and trust in the device, the results on usability are even better, with a SUS score of over 90, a RTLX score of 5.31 (SD: 11.38) and the fastest times for reading and writing keys. On the physical side, the sticker fared worse than the smartcard, with the main concern being that the storage and transmission of the keys being not that secure. Participants feared that the keys could just be read by an attacker being in close proximity to the sticker with a reading device, similar to contactless payment fraud. Additionally, while the sticker itself being easily transportable, participants had a feeling that the sticker could come loose easily and get lost.

The third place goes to the QR-Device, with the next highest usability score, the third lowest RTLX score and only slightly slower times for reading and writing keys than the smartcard. While loosing out on transportability and receiving criticism for being an extra device, as well as size and weight, it still achieved better results for securely storing the keys, fun while using and trust compared to the forth device, the USB-keychain.

The fourth place is taken by the USB-Keychain. Even though USB drives a common and nearly everybody already has experience with them, it does not mean they provide a great user experience. In the specific use case for this thesis, the USB rated second lowest on the SUS scale for usability, on the RTLX scale and on the timing measurements for storing and retrieving keys. This trend continues for perceived risk of use, trust and the overall HCTS trust score. The only aspect rated really positively was the transportability of the prototype and the number of times the prototype was criticized during the interview is fairly low.

The paper prototype places fifth, with bad SUS scores, a mid-high task load index, by far the longest time taken to retrieve the keys combined with low trust and a high perceived risk due to non existing security features and the key available in plaintext.

7 Discussion

This section provides explanations for some of the more special data points of the study, a short answer to the proposed research questions, tries to places the results in the context of a more realistic use case scenario and provide some ideas for possible future work. It also lists limitations of the study, lessons learned and possible improvements for the prototypes.

7.1 Noticeable Observations

This subsection provides some explanations for noticeable and interesting data points of the results.

Time difference USB task 1 and 2: A difference of around 5 seconds can be observed for the USB-Keychain between task 1 and task 2. While one would expect a time for both tasks that is quite similar, as the required user interaction is very similar, the difference can be explained by the used hardware for the workstation computers. Participants struggled way more to insert the USB drive into the side USB port of the surface, than with the thinkpads USB port, with one participant nearly taking half a minute until the key was saved on the USB.

Preference to a purely digital alternative: The level of agreement for the tangibility statement "I would prefer this device to a purely digital alternative." was generally rather low. This can be explained under the assumption that a purely digital key transmission would allow for simpler user interaction or even if not the case would at least not require the user to carry an additional device with them and fiddle with plugging it in or placing it on a reader or holding it in front of the camera. The low values for paper can be explained by the unnecessary complex interaction for entering the paper key, as in all previous statements too.

Secure storage: For the statement "I believe the device is keeping my keys secure", both the paper prototype and the USB-keychain received ratings worse, than the other devices. A possible explanation could be, that when looking at how the keys are stored on each device, for the USB and paper, the participants could either directly see the plaintext key on the paper, or the key file they saved and opened on the USB drive. The feeling of keeping the keys secure on the smartcard and the QR-Device, could be based on the prior experiences the participants already have with these prototypes. For the smartcard, similar experiences would be the use of banking cards, which we generally assume keep our data secure, an assumption that could easily be brought over to the smartcard prototype, as it follows the same user interactions of having to be inserted into a reader to be accessed. For the QR-Device, due to the simulation with a smartphone, the participants could have associated the QR-Device with their own smartphone, which often contains and keeps sensitive data secure. The larger amount of neither agree nor disagree for the sticker could be based on the absence of experience of using a RFID token to store data, and therefore not having a mental model on how that device works and is storing the keys.

Perceived Risk: The paper prototype and the USB-keychain both received high scores for perceived risk when using the devices. As this is somewhat persistent to the evaluations for secure storage, so a similar explanation can be used. The high perceived risk for paper can be explained by the key being printed in plaintext on the paper, which would allow easy access to the key in case the paper gets lost or stolen. The same can be said for the USB, as the key is only a file on the USB drive, an attacker could easily gain access to the key by physically accessing the USB and plugging it into any USB port. The highest lowest level of risk is reported for the smartcard, presumably because of the perceived secure transmission of keys, as well as the secure storage of the keys.

Benevolence: For the benevolence subscale, all devices received low scores, with the exception paper, which placed even lower. This could be caused by the way the individual statements of the subscale were worded. Consider the statement of HCTS item 5 ("I believe that the system will do its best to help me if I need help."): Most prototypes are from a user perspective simple single purpose devices, without even any option of a direct with the prototype. The one exception could be the QR-Device, as this is the only device that by itself reacts to user inputs and could be able to provide help or guidance if the user were lost or an error occurred while interacting with the prototype. For all other prototypes, such help would have to be provided via the accompanying software used to interact with the devices. As the software used in the user study was only a simple test application without any guidance or help on the use with the prototypes, the low ratings for benevolence could originate from there.

Competence: The values for competence do not differ much between the different prototypes, with the exception of paper. A possible explanation my be due to the way the statements are worded, the focus is on the single purpose of transporting cryptographic keys and with the devices not even having any additional functionalities, apart from getting keys on and off the devices, these ratings could be explained. If the devices had additional functionalities (as some improvements will be mentioned in chap. 6.4), the overall competence ratings would probably be higher, as the HCTS 9 statement is about the system having all expected functionalities for a key exchange device.

Trust: The ratings for trust could be assumed to be the inverse of the ratings for perceived risk, but this is not completely the case. Due to the way, the individual statements are worded, some room for interpretation is left when answering the questions, which in turn may explain why trust does not completely match the inverse of perceived risk. When comparing the median values, the smartcard has the highest trust (3.67) and the lowest perceived risk (2.33), followed by the RFID sticker with a median trust of 3.5 and median risk of 2.67. For the USB and QR-Device this does not hold up directly, as both have the same median trust of 3.33 but the QR-Device has a lower perceived risk with 3.0 compared to the USB 3.17.

Overall HCTS: The overall HCTS scores match the results of the statements on key storage and transmission, where USB and sticker also received worse ratings than the smartcard or the QR-Device.

Usage of prototypes: Sticker and smartcard have the highest potential usages in a work related environment, whether it is work-related data or email. While this could be caused by the high usability of both devices, it could also be caused by the fact, that many people already use smartcards in a work related setting. Two usage categories do stand out for the smartcard, online-banking and critical infrastructure. It seems the smartcard is the only prototype seen secure enough to store keys for highly sensitive data. overall the RFID sticker would be used for most categories, with the exception of high sensitivity ones, in contrast to paper whose usage is unexpectedly low, due to the general low usability of the paper prototype.

7.2 Answering the Research Questions

After evaluating all devices in regards to their usability, tangibility and user trust and security, a short summarizing answer on each research question is given.

The first question ”Q1: How do users perceive the usability of different key exchange devices (compared to a baseline)?”, can be directly answered by looking at the results of the questions that were specifically geared towards measuring usability. The proposed prototypes were perceived quite differently by the users, as only the RFID-Sticker manages to beat the smartcard baseline in terms of usability. The other two prototypes achieve a lower level of usability, with the QR-Device being better than the USB-Keychain. The paper baseline is beaten by all prototypes and therefore can only serve as a bad example of how an offline key exchange token should not look like.

For the second question ”Q2: How do the differences in tangibility affect the user experience?” the answer is not so clear at all. While specific questions towards the tangibility and form factor were asked, positive results here do not have to correlate with high ratings of usability. From the question we can discern the (somehow expected) result that smaller devices are easier transported, but may not all feel handy to the same degree. Size does also not directly correlate with the believe of easier loosing a device, as the fairly large paper and the second smallest (USB-Keychain) are both seen of being at the highest risk for getting lost. While holding a physical object in their hands gives the users a feeling of being in control, only the two prototypes with the simplest user interaction would be preferred by half the participants to a purely digital alternative.

While we are aware, that the answer to this question is quite vague, we acknowledge that these results may be caused by the lack of direct tangible interaction with the prototypes themselves. Just inserting the card into the reader or plugging an USB drive in, is not the height of possible tangible interaction, that these key storage devices could have, if e.g. compared to seedsigner cryptocurrency signings device, with its multiple buttons and selection joystick.

For the last question ”Q3: How do the different hardware devices effect the perception of trust and security in the devices?” large differences can be observed between the prototypes, e.g. only the smartcard and the QR-Device had more than half the participants agreement on keeping the keys secure. Lower differences between the prototypes were observed, when participants were asked about the security of key transmission, with the exception of paper all devices received similar results. It is interesting, that paper received the lowest scores for having a secure transmission even though it is the one method which is controlled most by the user, as the entry each character, representing some bit sequence of the key, is directly a result of the user pressing the corresponding key on the keyboard. These results can also be seen in the trust placed in each prototype. Here, the smartcard and the RFID-

7.3 Towards a more realistic use case scenario

This sections aims to provide an outlook and some ideas on what would constitute a more realistic use case scenario, as the one proposed in chap 3.1 is a simplified one, leaving out some important outside factors.

The first thing to mention here is, that our use case only covered the transport of a small amount of keys for individual users. If a larger amount of keys has to be transported, regardless of the source and destination location, the concepts presented in this paper can not be used anymore. This is caused directly by technical limitations of either the storage capacity being too small or the low bandwidth with which the keys are transmitted. Only the USB-keychain has the capabilities of storing a large amount of keys in addition to a high enough bandwidth to save and retrieve hundreds of keys and more.

Focusing back on the use case of only transporting a few keys at a time, for the way the keys were transported in the user study, only the sticker managed to beat the smartcard in usability but lost out on trust, perceived security and, due to participants fearing easy loss of the sticker, transportability.

Another aspect is related to the user interaction. This is the only aspect were the sticker managed beat the smartcard, but many modern smartcards also support contactless communication This in turn allows a smartcard to easily use the same simple user interaction concept of the sticker. Of course, this could lose the smartcard some of the perceived security and trust for the safe storage and transmission of the keys it currently has. This does not work the other way round, the sticker has no way to imitate the contact based communication of the smartcard, while still remaining its sticker unique properties. If the sticker would get the contact based communication like the smartcard and get rid of its sticker properties, the end result would again be something akin to a smartcard.

For a more realistic use case, consider that in many enterprise environments, which was what the presented use case tried to simulate, smartcards are already widely in use. Even though not primarily used as a key transport device, often employees have their private keys for signing and decrypting emails stored on a smartcard. While this could easily be done with a RFID-Sticker solution, this leaves the other purpose of the employee smartcards out. The more important aspects of the employee smartcard are, not the storing of cryptographic keys, but for access to the company grounds or for unlocking the employees workstation. While this could too be replicated with the sticker, the important aspect of quickly identifying authorising the employee to be allowed on company grounds can not be replicated. Normally, employees would have to wear their badges visible, so spotting unauthorized persons is fast an easy, as potentially any person without a visible badge is suspect. These consideration should be kept in mind, when trying to create a key transport device for individuals in an enterprise environment.

There is an other possible use case for storing and transporting cryptographic keys on an individual level, namely in the area of personal identification and authentication. For this use case, smartcards are already in use, in the form of national id cards. On the basis of a EU project for standardized electronic identification this allows EU/EEU citizens to use various public services, independent of their country of origin [46]. Considering the positive results of smartcards, this could be an interesting use case scenario or rather research object on the usability or trust for smartcard based digital identification methods, as from a practical standpoint it should not matter too much if the device transports symmetric keys for encryption or authentication and identification data instead.

When additionally taking into consideration, the rise of mobile payments worldwide [47] and the possibility to store drivers licences or state id in digital wallets on our phones [48], this could also be the future of personal ”offline” key storage, instead of using additional devices or smartcards.

7.4 Improving the prototypes

Taking the user feedback into account in this section some improvements for the prototypes are presented.

For paper a creative approach to increase the security of the prototype was suggested by P_{16}, namely the use of magic ink. Two possible solutions come to mind, either ink that is invisible unless under UV light or similar, to make the paper key appear less conspicuous or alternatively, ink that vanishes automatically after some time, rendering the paper unable to provide the key.

P_2 mentioned the use of RFID signal blocking technology to increase the security of the RFID-Sticker prototype. Depending on the exact implementation, this could significantly reduce the usability, if the object the sticker is attached to has to be taken out of a signal blocking case, even more so considering most participants attached the sticker to their smartphone. A possible option could be a slider, that only has to be slid open and can be closed again after the key has been transmitted.

Two direct improvements can be made to the USB prototype. Instead of having the user manually select the file, a more user friendly option could exist, if the file extension of the key file were directly associated with the key import application on the OS level, so the key could be directly imported just by trying to open it via the file browser. The other suggestion was to directly automate the user interaction, so the user only has to plug in the device and click ”save”. Everything else is then directly handled by the application.

A common complaint for the QR device the UI not being clear enough which of the saved keys on the devices is for which application. This could easily be implemented and could drastically increase the usability of the QR-Device.

7.5 Limitations

This study setup has some inherent limitations that should be kept in mind. The first one is based on the recruiting procedure. As the recruiting channels are geared towards the recruitment of students, the resulting group of participants tends to be a younger and higher educated subset of the general population. This could lead to a better evaluation of the prototypes, as younger people are often more open towards technology. The sample size is with a total of 16 participants quite small, which further reduces the validity when trying to generalize over all possible users.

Another limitation is, that the user interactions for each prototype are quite simplified, e.g. a card only has to be plugged in. This may lead to better usability scores than when using the prototypes in a more realistic setting. The next limitation is the absence of specific security features for the prototypes. While this decision was made consciously, it could lead to all devices being perceived as less secure. The comparability of the devices relative to each other should not be impacted, though.

8 Conclusion and Future Work

This chapter summarizes the master thesis and gives an outlook into further possible research on this topic.

8.1 Conclusion

With the rise of quantum computing and therefore the increasing risk of classical cryptosystems being broken, other methods for the exchange of cryptographic keys have to be explored.

One such method is offline key distribution, during which keys are not send over the internet or classic communication channels, but instead by sending or transporting a physical object containing the keys to the target location. This can take many forms, from sending hard drives via postal service to personal transport of keys printed on paper. As this area of physically transporting keys has not really been the focus of much research this far, this thesis aimed to provide some first insights into the usability, tangibility, perceived security and trust in devices that can be used for the key exchange.

For that a specific use case was developed and , with the help of a focus group, three concepts, for key exchange devices, created. These devices, being based on RFID stickers, regular keychain USB drives and a dedicated hardware device using QR codes to read and display keys, were then evaluated in a user study.

In comparison with two baseline devices of methods commonly used to transport keys, printed on paper and stored on a smartcard, only one of the concepts could convince the study participants in the aspect of usability and was rated better than the smartcard. Apart from the USB sticker, the smartcard was generally regarded as the best option for a key transport device, due to high levels of trust and security perceived by the participants when using that device, which can be explained as many participants already had experience with the use of smartcards, even though the specific use case was different. The other devices using QR codes or USB generally scored lower on all relevant metrics and personal evaluations. They also faced more criticism and negative feedback than the smartcard and sticker.

8.2 Future Work

Some aspects have not been answered or covered in this thesis, that warrant further study.

None of the prototypes had explicit security features, so it could be interesting to research users opinions on usability, security and trust for more complex devices, with security features that explicitly have to be interacted with, in comparison to the very simple key transport devices presented in this paper. This could also directly be combined with a few of the device improvements mentioned in chap 7.4.

In the same vein would be to measure users attitudes towards the devices when faced with more complex tasks that goes further than just plugging the device in and clicking the ”transfer key” button. This could work well with a longer time frame for the study, in which the participants have to use the devices daily to perform a specified task, preferably at different locations, so the participants have to transport the device in their day to day life.

Another potential avenue for further work could be, to not use dedicated hardware devices for the key storage and transport, but instead integrate the key storage in wearable devices. This could positively affect fear of loosing the device and could lead to interesting user interaction concepts.

As described in chap. 7.3, the future potential in moving towards app/digital wallet based solutions could also be a useful topic for additional research. While this is, at best, only adjacent to offline key distribution, it could still provide insights into the use of personal smartphones as key transport devices.

Inhalt der beigelegten CD

Da die Abgabe digital erfolgt liegt keine CD bei. Im Rahmen der digitalen Abgabe wird folgendes mit eingereicht:

1. Quellcode aller Programme (QR-Device app, Desktop App, arduino Kartenleser)
2. Transkript und Auswertung des Fokusgruppen Interviews
3. Zusammenschnitt aller Videoaufzeichnungen der Studiendurchläufe
4. Aus den Videos extrahierte Audioaufzeichnungen
5. Audioaufzeichnungen der Schlussinterviews
6. Transkript der Interviews
7. Latex Quellcode der Ausarbeitung
8. Rohdaten export aller validen Fragebögen der Nutzerstudie
9. Timelogs der beiden Workstations und des QR-Devices
10. Jupyter Datenauswertungsskript

References

Literature

[1] Claudia Ziegler Acemyan et al. “2FA Might Be Secure, But It’s Not Usable: A Summative Usability Assessment of Google’s Two-factor Authentication (2FA) Methods”. en. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62.1 (Sept. 2018), pp. 1141–1145. ISSN: 2169-5067, 1071-1813. DOI: 10.1177/1541931218621262. URL: http://journals.sagepub.com/doi/10.1177/1541931218621262 (visited on 01/05/2022).

[2] Romain Alléaume et al. “Quantum key distribution and cryptography: a survey”. In: Dagstuhl Seminar Proceedings (DagSemProc) 9311 (2010). Ed. by Samual L. Braunstein et al., pp. 1–29. ISSN: 1862-4405. DOI: 10.4230/DagSemProc.09311.3. URL: https://drops.dagstuhl.de/opus/volltexte/2010/2361.

[3] Wei Bai et al. “An Inconvenient Trust: User Attitudes toward Security and Usability Tradeoffs for Key-Directory Encryption Systems”. In: *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. Denver, CO: USENIX Association, June 2016, pp. 113–130. ISBN: 978-1-931971-31-7. URL: https://www.usenix.org/conference/soups2016/technical-sessions/presentation/bai.

[4] Daniel V. Bailey and Horst Görtz. ““ Typing ” passwords with voice recognition : How to authenticate to Google Glass”. In: 2014.

[5] Aaron Bangor, Philip Kortum, and James Miller. “Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale”. In: *J. Usability Studies* 4.3 (May 2009), pp. 114–123.

[6] Deborah Barreau and Bonnie A. Nardi. “Finding and reminding: file organization from the desktop”. en. In: *ACM SIGCHI Bulletin* 27.3 (July 1995), pp. 39–43. ISSN: 0736-6906. DOI: 10.1145/221296.221307. URL: https://dl.acm.org/doi/10.1145/221296.221307 (visited on 09/04/2022).

[7] Charles H Bennett et al. “Generalized Privacy Amplification”. en. In: *IEEE TRANSACTIONS ON INFORMATION THEORY* 41.6 (Nov. 1995), p. 9.

[8] Christina Braz and Jean-Marc Robert. “Security and usability: the case of the user authentication methods”. en. In: *Proceedings of the 18th international conference on Association Francophone d’Interaction Homme-Machine - IHM ’06*. Montreal, Canada: ACM Press, 2006, pp. 199–203. ISBN: 978-1-59593-350-8. DOI: 10.1145/1132736.1132768. URL: http://portal.acm.org/citation.cfm?doid=1132736.1132768 (visited on 09/04/2022).

[9] John Brooke. *"SUS-A quick and dirty usability scale." Usability evaluation in industry*. ISBN: 9780748404605. CRC Press, June 1996. URL: https://www.crcpress.com/product/isbn/9780748404605.

[10] Stéphane Ciolino, Simon Parkin, and Paul Dunphy. “Of Two Minds about Two-Factor: Understanding Everyday FIDO U2F Usability through Device Comparison and Experience Sampling”. In: *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security*. SOUPS’19. Santa Clara, CA, USA: USENIX Association, 2019, pp. 339–356. ISBN: 9781939133052.

[11] Lorrie Cranor and Simson Garfinkel. *Security and Usability*. O’Reilly Media, Inc., 2005. ISBN: 0596008279.

[12] Sanchari Das et al. “A qualitative study on usability and acceptability of Yubico security key”. en. In: *Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust - STAST ’17*. Orlando, Florida: ACM Press, 2018, pp. 28–39. ISBN: 978-1-4503-6357-0. DOI: 10.1145/3167996.3167997. URL: http://dl.acm.org/citation.cfm?doid=3167996.3167997 (visited on 01/05/2022).

[13] Mehrdad Dianati et al. “Architecture and protocols of the future European quantum key distribution network”. en. In: *Security and Communication Networks* 1.1 (Jan. 2008), pp. 57–74. ISSN: 19390114, 19390122. DOI: 10.1002/sec.13. URL: https://onlinelibrary.wiley.com/doi/10.1002/sec.13 (visited on 02/13/2022).

[14] Serge Egelman and Eyal Peer. “Scaling the Security Wall: Developing a Security Behavior Intentions Scale (SeBIS)”. en. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul Republic of Korea: ACM, Apr. 2015, pp. 2873–2882. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702249. URL: https://dl.acm.org/doi/10.1145/2702123.2702249 (visited on 08/29/2022).

[15] Cori Faklaris, Laura Dabbish, and Jason I. Hong. “A Self-Report Measure of End-User Security Attitudes (SA-6)”. In: *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security*. SOUPS’19. Santa Clara, CA, USA: USENIX Association, 2019, pp. 61–77. ISBN: 9781939133052.

[16] Thomas Franke, Christiane Attig, and Daniel Wessel. “A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale”. In: *International Journal of Human–Computer Interaction* 35.6 (2019), pp. 456–467. DOI: 10.1080/10447318.2018.1456150. eprint: https://doi.org/10.1080/10447318.2018.1456150. URL: https://doi.org/10.1080/10447318.2018.1456150.

[17] Sanam Ghorbani Lyastani et al. “Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication”. en. In: *2020 IEEE Symposium on Security and Privacy (SP)*. San Francisco, CA, USA: IEEE, May 2020, pp. 268–285. ISBN: 978-1-72813-497-0. DOI: 10.1109/SP40000.2020.00047. URL: https://ieeexplore.ieee.org/document/9152694/ (visited on 01/05/2022).

[18] Intenso International GmbH. *MICRO LINE USB 2.0*.

[19] Siddharth Gulati, Sonia Sousa, and David Lamas. “Design, development and evaluation of a human-computer trust scale”. en. In: *Behaviour & Information Technology* 38.10 (Oct. 2019), pp. 1004–1015. ISSN: 0144-929X, 1362-3001. DOI: 10.1080/0144929X.2019.1656779. URL: https://www.tandfonline.com/doi/full/10.1080/0144929X.2019.1656779 (visited on 08/29/2022).

[20] Amina Harit, Abdellah Ezzati, and Rachid Elharti. “Internet of things security: challenges and perspectives”. en. In: *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*. Cambridge United Kingdom: ACM, Mar. 2017, pp. 1–8. ISBN: 978-1-4503-4774-7. DOI: 10.1145/3018896.3056784. URL: https://dl.acm.org/doi/10.1145/3018896.3056784 (visited on 02/14/2022).

[21] Sandra G. Hart and Lowell E. Staveland. “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research”. en. In: *Advances in Psychology*. Vol. 52. Elsevier, 1988, pp. 139–183. ISBN: 978-0-444-70388-0. DOI: 10.1016/S0166-4115(08)62386-9. URL: https://linkinghub.elsevier.com/retrieve/pii/S0166411508623869 (visited on 09/01/2022).

[22] Kat Krol et al. “"They brought in the horrible key ring thing!" Analysing the Usability of Two-Factor Authentication in UK Online Banking”. en. In: *Proceedings 2015 Workshop on Usable Security*. San Diego, CA: Internet Society, 2015. ISBN: 978-1-891562-40-2. DOI: 10.14722/usec.2015.23001. URL: https://www.ndss-symposium.org/ndss2015/ndss-2015-usec-programme/they-brought-horrible-key-ring-thing-analysing-usability-two-factor-authentication-uk-online (visited on 02/14/2022).

[23] Richard Krueger. *Designing and Conducting Focus Group Interviews*. en. URL: https://www.eiu.edu/ihec/Krueger-FocusGroupInterviews.pdf (visited on 02/14/2022).

[24] Jonathan Lazar. *Research methods in human computer interaction*. en. 2nd edition. Cambridge, MA: Elsevier, 2017. ISBN: 978-0-12-805390-4.

[25] Igor Lopez and Marina Aguado. “Cyber security analysis of the European train control system”. en. In: *IEEE Communications Magazine* 53.10 (Oct. 2015), pp. 110–116. ISSN: 0163-6804. DOI: 10.1109/MCOM.2015.7295471. URL: http://ieeexplore.ieee.org/document/7295471/ (visited on 08/29/2022).

[26] D. Harrison Mcknight et al. “Trust in a Specific Technology: An Investigation of Its Components and Measures”. In: *ACM Trans. Manage. Inf. Syst.* 2.2 (July 2011). ISSN: 2158-656X. DOI: 10.1145/1985347.1985353. URL: https://doi.org/10.1145/1985347.1985353.

[27] Ralph C. Merkle. “Secure communications over insecure channels”. en. In: *Communications of the ACM* 21.4 (Apr. 1978), pp. 294–299. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/359460.359473. URL: https://dl.acm.org/doi/10.1145/359460.359473 (visited on 01/05/2022).

[28] Emile Morse et al. *Usability of PIV Smartcards for Logical Access*. en. Tech. rep. NIST IR 7867. National Institute of Standards and Technology, June 2012, NIST IR 7867. DOI: 10.6028/NIST.IR.7867. URL: https://nvlpubs.nist.gov/nistpubs/ir/2012/NIST.IR.7867.pdf (visited on 08/29/2022).

[29] Douglas O’Shaughnessy. “Invited paper: Automatic speech recognition: History, methods and challenges”. en. In: *Pattern Recognition* 41.10 (Oct. 2008), pp. 2965–2979. ISSN: 00313203. DOI: 10.1016/j.patcog.2008.05.008. URL: https://linkinghub.elsevier.com/retrieve/pii/S0031320308001799 (visited on 02/14/2022).

[30] Kseniia Palin et al. “How Do People Type on Mobile Devices? Observations from a Study with 37,000 Volunteers”. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’19. Taipei, Taiwan: Association for Computing Machinery, 2019. ISBN: 9781450368254. DOI: 10.1145/3338286.3340120. URL: https://doi.org/10.1145/3338286.3340120.

[31] Urbano B. Patayon and Nerico L. Mingoc. “Operating Systems Usability: A Comparative Study”. en. In: *JPAIR Multidisciplinary Research* 36.1 (Mar. 2019), pp. 92–103. ISSN: 2244-0445, 2012-3981. DOI: 10.7719/jpair.v36i1.683. URL: http://philair.ph/index.php/jpair/article/view/683 (visited on 09/04/2022).

[32] Joann Peck and Terry L. Childers. “Individual Differences in Haptic Information Processing: The “Need for Touch” Scale”. en. In: *Journal of Consumer Research* 30.3 (Dec. 2003), pp. 430–442. ISSN: 0093-5301, 1537-5277. DOI: 10.1086/378619. URL: https://academic.oup.com/jcr/article-lookup/doi/10.1086/378619 (visited on 08/29/2022).

[33] W. Rankl. *Smart card handbook*. en. 4th ed. Chichester, West Sussex, U.K: Wiley, 2010. ISBN: 978-0-470-74367-6.

[34] Ken Reese et al. “A Usability Study of Five Two-Factor Authentication Methods”. In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 357–370. ISBN: 978-1-939133-05-2. URL: https://www.usenix.org/conference/soups2019/presentation/reese.

[35] Sarah Delgado Rodriguez et al. “PriKey – Investigating Tangible Privacy Control for Smart Home Inhabitants and Visitors ”. In: NordiCHI ’22 (2022). delgado2022nordichi. URL: http://www.florian-alt.org/unibw/wp-content/publications/delgado2022nordichi.pdf.

[36] David E. Rumelhart and Donald A. Norman. “Simulating a Skilled Typist: A Study of Skilled Cognitive-Motor Performance”. In: *Cogn. Sci.* 6 (1982), pp. 1–36.

[37] Valerio Scarani and Christian Kurtsiefer. “The black paper of quantum cryptography: Real implementation problems”. en. In: *Theoretical Computer Science* 560 (Dec. 2014), pp. 27–32. ISSN: 03043975. DOI: 10.1016/j.tcs.2014.09.015. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304397514006938 (visited on 01/05/2022).

[38] Eva-Maria Schomakers et al. “Internet users’ perceptions of information sensitivity – insights from Germany”. en. In: *International Journal of Information Management* 46 (June 2019), pp. 142–150. ISSN: 02684012. DOI: 10.1016/j.ijinfomgt.2018.11.018. URL: https://linkinghub.elsevier.com/retrieve/pii/S0268401218307692 (visited on 09/01/2022).

[39] NXP Semiconductors. *MF1S503x - MIFARE Classic 1K - Mainstream contactless smart card IC for fast and easy solution development*. 194031. Rev. 3.1. Feb. 21, 2011.

[40] NXP Semiconductors. *NTAG213/215/216 - NFC Forum Type 2 Tag compliant IC with 144/504/888 bytes user memory*. 265332. Rev. 3.2. June 2, 2015.

[41] NXP Semiconductors. *MFRC522 - Standard performance MIFARE and NTAG frontend*. 112139. Rev. 3.9. Apr. 27, 2016.

[42] Frank Stajano. “Pico: No More Passwords!” en. In: *Security Protocols XIX*. Ed. by Bruce Christianson et al. Vol. 7114. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 49–81. ISBN: 978-3-642-25866-4 978-3-642-25867-1. DOI: 10.1007/978-3-642-25867-1_6. URL: http://link.springer.com/10.1007/978-3-642-25867-1_6 (visited on 01/17/2022).

[43] Jinyue Xia and Yongge Wang. “Secure Key Distribution for the Smart Grid”. en. In: *IEEE Transactions on Smart Grid* 3.3 (Sept. 2012), pp. 1437–1443. ISSN: 1949-3053, 1949-3061. DOI: 10.1109/TSG.2012.2199141. URL: http://ieeexplore.ieee.org/document/6205351/ (visited on 01/17/2022).

[44] Yulong Yang, Janne Lindqvist, and Antti Oulasvirta. “Text Entry Method Affects Password Security”. In: *CoRR* abs/1403.1910 (2014). arXiv: 1403.1910. URL: http://arxiv.org/abs/1403.1910.

Web-references

[45] Monica Chin. *File not found. A generation that grew up with Google is forcing professors to rethink their lesson plans*. Sept. 2021. URL: https://www.theverge.com/22684730/students-file-folder-directory-structure-education-gen-z.

[46] European Comission. *European Digital Identity*. URL: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en.

[47] David Curry. *Mobile Payments App Revenue and Usage Statistics (2022)*. May 2022. URL: https://www.businessofapps.com/data/mobile-payments-app-market/.

[48] Apple Press Release. *Apple announces first states signed up to adopt driver's licenses and state IDs in Apple Wallet*. Sept. 2021. URL: https://www.apple.com/newsroom/2021/09/apple-announces-first-states-to-adopt-drivers-licenses-and-state-ids-in-wallet/.